



# Paramétrisation des modèles physiologiques toxico/pharmacocinétiques

Céline Brochot

## ► To cite this version:

Céline Brochot. Paramétrisation des modèles physiologiques toxico/pharmacocinétiques. Sciences du Vivant [q-bio]. Université Pierre et Marie Curie - Paris VI, 2004. Français. NNT : . tel-00107374

**HAL Id: tel-00107374**

**<https://theses.hal.science/tel-00107374>**

Submitted on 18 Oct 2006

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITE PIERRE ET MARIE CURIE - PARIS 6

**THESE**

Pour obtenir le titre de

**DOCTEUR DE L'UNIVERSITE DE PARIS 6**

Spécialité Biomathématiques

Ecole doctorale : Santé Publique et Sciences de l'Information Biomédicale

Présentée par

**Céline BROCHOT**

**PARAMETRISATION DES MODELES  
PHYSIOLOGIQUES TOXICO/PHARMACOCINETIQUES**

**Jury**

Pr. Guy THOMAS

Pr. Eric PARENT

Dr. Randall THOMAS

Dr. Frédéric Y. BOIS

Pr. Charles-André Cuenod

Président

Rapporteur

Rapporteur

Directeur de thèse

Examineur

## Remerciements

Je tiens tout d'abord à remercier sincèrement Monsieur le Docteur Frédéric Bois de m'avoir donné l'opportunité d'effectuer ma thèse à l'INERIS au sein de l'unité Toxicologie Expérimentale. Merci de m'avoir conseillée et orientée vers le DEA de biomathématiques, et ensuite de m'avoir fait découvrir la pharmacocinétique, les statistiques bayésiennes, sans oublier *MCSim*, bien sûr. Pour finir, merci d'avoir monter toutes ces collaborations pour me permettre de mener à bien mon projet.

Je suis très reconnaissante à Monsieur le Professeur Guy Thomas de me faire l'honneur d'accepter la présidence de mon jury de thèse.

Je tiens à remercier vivement Monsieur le Professeur Eric Parent et Monsieur le Professeur Randall Thomas d'avoir accepter d'examiner mes travaux de thèse, et d'en être les rapporteurs.

Je remercie Monsieur le Professeur Charles-André Cuénod de sa participation à mon jury de thèse, ainsi que pour la qualité de notre collaboration. Je lui suis reconnaissante de m'avoir invitée lors d'une séance d'imagerie par résonance magnétique. Merci à l'unité INSERM U494, et en particulier à Daniel Balvay, qui a pris le temps de m'expliquer ces techniques. Je tiens également à remercier toutes les personnes qui ont participé à la collecte des données.

Je remercie Monsieur le Professeur Janos Tóth de m'avoir aidé à comprendre les techniques d'agrégation et le logiciel *Mathematica*, et de m'avoir accueillie dans son laboratoire de mathématiques à l'Université de Budapest. En espérant que nos relations franco-hongroises repartent sur de meilleures bases ...

Je tiens également à remercier Madame le Docteur Marie-Christine Jaulent, ma tutrice de thèse, pour m'avoir écoutée et motivée pendant une période de doutes, et Madame le Docteur Sophie Lesieur pour avoir répondu à mes nombreuses questions administratives sur le déroulement d'une thèse.

Un merci amical à Sandrine, ma colocataire de bureau, pour son aide et son soutien. En espérant t'avoir rendu la pareille. Merci à l'équipe des modélisateurs / statisticiens (Billy, Cheikh, Sandrine). Merci à l'unité Toxicologie Expérimentale, et plus particulièrement à Aleksandra, Elsa, Simon, et Anne-Sophie pour les bons moments. Pour finir, merci à tous mes relecteurs.

Un merci chaleureux à ma famille et à la famille Cazalet Dance pour m'avoir toujours soutenue. Last but not least, merci à Rémi pour son aide (morale, informatique, et « orthographique ») tout au long de ces trois années. Une dernière pensée à notre petit bout qui a dû supporter tout le stress de fin de thèse dès les premiers instants de sa vie ...

# Table des matières

<b>1. Résumés</b>	<b>5</b>
1.1. Résumé	5
1.2. Abstract	6
<b>2. Présentation du sujet</b>	<b>7</b>
<b>3. Contexte et objectifs spécifiques</b>	<b>10</b>
<b>3.1. Les modèles physiologiques toxico/pharmacocinétiques</b>	<b>10</b>
3.1.1. Les processus ADME	11
A. Absorption	11
B. Distribution	12
C. Métabolisme	13
D. Excrétion	14
3.1.2. Spécification d'un modèle PBPK	14
3.1.3. Paramétrisation d'un modèle PBPK	17
3.1.4. Un outil pour l'évaluation de risque et les études thérapeutiques	18
A. Transpositions entre conditions expérimentales	19
B. Transpositions inter-espèce, inter- et intra-individu	19
<b>3.2. Calibration et analyse bayésienne</b>	<b>21</b>
3.2.1. L'inférence bayésienne	21
3.2.2. Modèles statistiques	22
A. Distributions <i>a priori</i>	22
B. Vraisemblance	22
3.2.3. Mise en œuvre de la calibration bayésienne	23
A. Méthodes Monte Carlo par chaînes de Markov	23
B. L'algorithme de Metropolis-Hastings	24
C. Contrôle de la convergence	25
D. Logiciels	26
<b>3.3. L'imagerie par résonance magnétique</b>	<b>27</b>
3.3.1. Rappels de physique	27
A. Champ et moment magnétiques	27
B. Le principe de résonance	28
C. Propriétés magnétiques de la matière	28
3.3.2. Le signal de résonance magnétique nucléaire	29
A. Action d'un champ magnétique sur une population d'atomes	29

Table des matières	2
B. Le signal RMN et ses caractéristiques	30
C. Le phénomène de relaxation	32
3.3.3. Obtention de l'image de résonance magnétique nucléaire	33
A. Codage de la position du voxel	33
B. Gradient de champ magnétique	34
3.3.4. Agents de contraste	34
3.3.5. La place de l'IRM dans la recherche pharmaceutique	34
<b>3.4. L'agrégation de système d'équations différentielles</b>	<b>36</b>
3.4.1. L'agrégation linéaire	36
A. Agrégation linéaire exacte	36
B. Agrégation linéaire contrainte : la méthode DCAL	38
3.4.2. Exemple d'agrégation exacte d'un système de dimension 3	39
A. Définition du système	40
B. Les schémas d'agrégation	40
C. Formulation du système agrégé	41
<b>3.5. Objectifs spécifiques</b>	<b>42</b>
<b>4. L'analyse statistique bayésienne de données toxicocinétiques</b>	<b>44</b>
<b>L'Analyse Statistique Bayésienne de Données Toxicocinétiques</b>	<b>45</b>
RESUME	46
SUMMARY	47
INTRODUCTION	48
QUESTION TOXICOLOGIQUE ET DONNEES EXPERIMENTALES	50
LE PROCESSUS D'ANALYSE DES DONNEES	52
CONCLUSION	66
ANNEXE : EQUATIONS DEFINISSANT LE MODELE A 2-COMPATIMENTS.	68
REFERENCES	68
<b>5. Modélisation physiologique de données spatio-temporelles d'imagerie médicale</b>	<b>71</b>
<b>Bayesian physiological pharmacokinetic modeling of tumors' microcirculation magnetic resonance imaging data: Application to the evaluation of an antiangiogenic treatment.</b>	<b>72</b>
RESUME	73

Table des matières	3
ABSTRACT	74
INTRODUCTION	75
METHODS	75
RESULTS	83
DISCUSSION	88
REFERENCES	90

## **6. Utilisation de traceurs non-toxiques dans les études toxicocinétiques 93**

### **Use of a chemical probe to increase safety for human volunteers in toxicokinetic studies 94**

RESUME	95
ABSTRACT	96
INTRODUCTION	97
METHODS	99
RESULTS	107
DISCUSSION	113
REFERENCES	115

### **Lumping in pharmacokinetics 118**

RESUME	119
ABSTRACT	120
INTRODUCTION	121
METHODS	122
APPLICATIONS OF MATHEMATICAL LUMPING IN THE PK FIELD	125
DISCUSSION	132
REFERENCES	134

<b>7. Conclusion et perspectives</b>	<b>136</b>
<b>8. Abréviations</b>	<b>140</b>
<b>9. Références</b>	<b>141</b>
<b>10. Publications, communications, et collaborations</b>	<b>155</b>
10.1. Publications	155
10.2. Communications	155
10.3. Collaborations	155



# 1. Résumés

## 1.1. Résumé

Les modèles physiologiques toxicocinétiques (PBPK pour Physiologically-Based Pharmacokinetic) permettent de décrire la biodistribution d'une substance (c'est-à-dire, son administration, distribution, métabolisme, et excrétion) au sein d'un organisme. Cependant, l'ajustement d'un modèle PBPK à des données expérimentales peut se révéler difficile en raison de la complexité du modèle, du nombre important de paramètres à estimer, et de la nature des données toxicocinétiques (généralement peu nombreuses, éparses et entachées d'erreur de mesure). Pour pallier à ces difficultés, nous proposons de développer deux protocoles expérimentaux facilitant la paramétrisation des modèles PBPK, par l'apport d'informations complémentaires sur les paramètres des modèles : la collecte de données sur la distribution spatio-temporelle d'une substance dans l'organisme par des techniques d'imagerie, et l'utilisation de traceurs non-toxiques lors d'études toxicocinétiques.

Pour notre premier protocole, nous avons appliqué l'analyse d'image au développement d'un modèle PBPK décrivant la biodistribution d'un agent de contraste, Vistarem<sup>®</sup>. Grâce à la modélisation des données recueillies par imagerie par résonance magnétique, nous avons pu caractériser la vascularisation du corps entier et celle d'une tumeur greffée, ainsi qu'évaluer statistiquement les effets d'un traitement antiangiogénique.

Ensuite, nous avons introduit l'utilisation de traceurs explorateurs lors d'études toxicocinétiques. Nous avons développé des protocoles d'exposition à une substance toxique et à un traceur non-toxique pour les humains. Sur la base d'expositions simulées, nous avons montré qu'en administrant aux sujets deux substances, il est possible de réduire la dose d'exposition du toxique (par 4 dans notre exemple d'application) tout en maintenant la précision sur l'estimation des paramètres d'intérêt. L'inconvénient majeur de ce protocole d'exposition est la lourdeur des calculs induits (la calibration de deux modèles PBPK est nécessaire). Afin de réduire la complexité des calculs engendrés, nous avons proposé d'appliquer des techniques d'agrégation de systèmes d'équations différentielles. L'essentiel de ces techniques est de réduire un système tout en conservant sa dynamique. Sur des exemples simples, nous avons montré leur potentiel en toxico/pharmacocinétique.

Chacun des deux protocoles développés dans ce travail inclut, en plus de la substance d'intérêt, l'utilisation d'une substance "annexe". Notre travail montre que l'utilisation de cette substance annexe permet de renforcer la connaissance sur l'anatomie et la physiologie du sujet considéré. Ceci conduit donc à une meilleure détermination de l'action de la substance d'intérêt. Une suite logique de notre travail serait de coupler ces deux protocoles, c'est-à-dire de développer des protocoles d'exposition à une substance toxique et à une substance visualisable par une technique d'imagerie quantitative.

Mots-Clés : modèles PBPK; analyse statistique bayésienne; imagerie par résonance magnétique; traceur; méthode d'agrégation.

## 1.2. Abstract

Physiologically-based pharmacokinetic (PBPK) models are used to describe the biodistribution (administration, distribution, metabolism and excretion) of substances in the body. However, fitting a PBPK model to experimental data can be difficult because of the model complexity, the large number of parameters, and the amount of toxicokinetic data available (typically, few and sparse). To overcome these difficulties, we propose to develop two experimental protocols to parameterize PBPK models: the gathering of data on the spatio-temporal distribution of a substance in the body by imaging techniques, and the use of non-toxic tracer (or probe) for toxicokinetic studies. These two protocols should bring complementary information for the model parameters.

For the first protocol, we applied image analysis to the development of a PBPK model for a contrast agent, Vistarem<sup>®</sup>. Thanks to the modeling of the experimental data obtained with magnetic resonance imaging, we have characterized the whole-body vascularization, and the one of a grafted tumor. Moreover, we evaluated statistically the effects of an antiangiogenic treatment.

We also introduced the use of probes in toxicokinetic studies. We developed exposure scenarios to a chemical and to a probe for humans. With simulated data, we showed that it is possible to reduce the toxic exposure dose (by 4 in our example) while maintaining precision in parameter estimates, if the toxic exposure is completed with an exposure to a probe. The main disadvantage of this protocol is the complexity of the data analysis (two PBPK model calibrations should be performed). To reduce this complexity, we proposed to apply lumping methods. Such methods aim at reducing differential equation systems. With simple examples, we showed their usefulness for toxico/pharmacokinetic models.

Each protocol developed here includes a supplementary substance. Our work shows that the use of this substance reinforces knowledge on the anatomy and physiology of the subject studied. This yields to a better determination of the action of the substance of interest. A logical following of our work is to couple the two protocols, that is to say to develop exposure scenarios to a toxicant and to a substance that can be followed by an imaging technique.

Keywords: PBPK models; Bayesian analysis; magnetic resonance imaging; probe; lumping methods.

## 2. Présentation du sujet

Quantifier les effets, néfastes ou bénéfiques, découlant de l'exposition à une substance chimique est une problématique majeure en toxicologie et pharmacologie. En effet, il est essentiel de pouvoir prédire les risques encourus par les populations exposées à un agent toxique, de même que les bénéfices attendus suite à la prise d'une substance thérapeutique.

Pour protéger les populations, les toxicologues établissent des indices toxicologiques correspondant en général à des seuils au-delà desquels des effets se produisent. Comme exemple d'indice toxicologique, on peut citer la plus faible dose d'exposition pour laquelle aucun effet n'a été observé (cet indice est plus connu sous le sigle NOAEL, pour no-observed-adverse-effect level). Ces indices sont typiquement déterminés à l'aide de données obtenues à partir d'expérimentations animales, et sont directement spécifiques de la substance étudiée, de l'effet et du scénario d'exposition. Une fois l'indice défini pour une espèce animale, différents facteurs d'incertitude sont ensuite appliqués pour le transposer à une autre espèce (typiquement 10 pour une transposition animal/humain), ou encore pour tenir compte de la variabilité au sein d'une même population, etc. La valeur d'un facteur d'incertitude n'a cependant pas de réel fondement biologique, et est souvent discutable (Beck and Clewell, 2001; Margni *et al.*, 2002). Pour limiter ces nombreuses incertitudes, des approches nouvelles, telles que les modèles dose-réponse fondés biologiquement et les modèles toxicocinétiques, sont maintenant préconisées pour l'évaluation des risques (Bruckner *et al.*, 2004; Sweeney *et al.*, 2001; US Environmental Protection Agency, 1996; Zeise *et al.*, 2002).

La problématique se pose différemment pour la recherche en pharmacologie, et en particulier dans l'industrie pharmaceutique. En effet, l'utilisation de modèles ayant une réalité biologique fait déjà partie intégrante du processus de développement des substances thérapeutiques (Balant and Gex-Fabry, 2000). Une des problématiques des industries pharmaceutiques est de sélectionner le plus tôt possible les molécules ayant un fort potentiel, afin notamment de réduire les coûts financiers inhérents à leurs développements (Bertrand *et al.*, 2000). En se basant sur la structure de la molécule, il est possible de prédire certains de ses paramètres pharmacocinétiques, tels que son volume de distribution (Ghafourian *et al.*, 2004), à l'aide de formules mathématiques appelées relations quantitatives structure-activité (ou QSAR pour Quantitative Structure-Activity Relationship) (Hansch *et al.*, 2004; Poulin *et al.*, 2001). Un des axes de recherche est de prédire le comportement pharmacocinétique des molécules candidates en intégrant ces données obtenues *in silico*, et/ou des données *in vitro* dans un modèle (Reigner *et al.*, 1997; Suzuki *et al.*, 1995; Theil *et al.*, 2003).

La toxicocinétique (ou pharmacocinétique pour les substances thérapeutiques) étudie qualitativement et quantitativement le devenir des substances toxiques dans l'organisme (biodistribution), dans le but de comprendre et de contrôler leurs effets. (Les termes toxicocinétique et pharmacocinétique seront utilisés dans la suite.) Les effets relatifs à l'exposition à une substance ne sont pas uniquement liés à la dose reçue lors de l'exposition, mais dépendent de la quantité de produit actif (parent ou métabolite) atteignant les cellules

cibles (Andersen, 1995a). De par la complexité des processus de biodistribution mis en jeu, la relation liant la dose efficace à celle d'exposition n'est pas forcément linéaire. Outre les caractéristiques physiologiques de l'organisme exposé, la distribution de la substance peut être influencée par d'autres facteurs, tels que l'âge, le sexe ou le niveau d'activité physique (Rowland and Tozer, 1989). La modélisation de la biodistribution de la substance dans l'organisme est appelée modélisation toxicocinétique.

Deux approches de modélisation pharmaco/toxicocinétique sont principalement utilisées : les modèles compartimentaux de type classique (Gibaldi and Perrier, 1982; Jacquez, 1999; Jacquez and Simon, 1993) et les modèles physiologiques (Balant and Gex-Fabry, 1990; Gerlowski and Jain, 1983; Nestorov, 2003). Ces modèles se basent sur la subdivision de l'organisme en compartiments, dans lequel la substance est supposée se distribuer uniformément. Seuls quelques compartiments liés par des échanges de substance définissent un modèle toxicocinétique classique. Typiquement, un compartiment central est défini, et des compartiments périphériques sont ajoutés jusqu'à l'obtention d'un ajustement correct du modèle aux données (un ou deux sont généralement nécessaires). Les compartiments ne sont donc pas nécessairement biologiquement interprétables, ni les paramètres du modèle. Ces modèles permettent de décrire de façon satisfaisante certaines cinétiques avec un minimum de paramètres (Rowland and Tozer, 1989).

Les modèles toxicocinétiques physiologiques (ou PBPK pour Physiologically Based Pharmacokinetic) peuvent décrire de façon plus détaillée les mécanismes connus de l'absorption, de distribution, du métabolisme et de l'excrétion. Contrairement aux modèles classiques, les compartiments et paramètres de ces modèles ont pour la plupart une signification physiologique. Les échanges de matière sont régis par les flux sanguins, lymphatiques, ou autres transports et diffusions identifiables physiologiquement. Ces modèles peuvent avoir différents niveaux de complexité, selon les buts suivis par la modélisation et les caractéristiques de la substance étudiée.

Les avantages des modèles toxicocinétiques classiques sont la simplicité de leur structure et leur rapidité d'ajustement. De par leur description paramétrique de l'organisme, les modèles PBPK sont plus satisfaisants pour comprendre les variabilités inter ou intra-individu, ou extrapoler et transposer les résultats à d'autres conditions expérimentales (inter-dose, inter-voie d'administration, inter-espèce ou encore inter-individu). Dans ce travail, nous nous intéresserons au développement de modèles PBPK et à leur ajustement à des données expérimentales.

Plusieurs méthodes ont été envisagées pour ajuster un modèle PBPK à des données expérimentales. La première consiste à fixer certains paramètres du modèle à l'aide de valeurs référencées dans la littérature, et d'estimer uniquement les paramètres d'intérêt (les constantes d'élimination de la substance, par exemple). Cette méthode a cependant des limites puisque les valeurs typiquement référencées sont des valeurs standards obtenues pour des hommes de type Caucasien, et ne reflètent donc pas l'hétérogénéité de ces valeurs au sein d'une population. Alternativement, Gelman *et al.* (1996b) ont proposé d'ajuster les modèles PBPK dans un cadre bayésien. Contrairement à la première méthode, une valeur unique n'est pas assignée à chaque paramètre mais une distribution de probabilité, reflétant la connaissance actuelle (avec son incertitude) sur le paramètre et sa variabilité au sein de la population étudiée (Nestorov, 2001). Il est possible

de différencier l'incertitude et la variabilité en utilisant un modèle de population (Bois, 2001; Sheiner and Ludden, 1992; Wakefield, 2000). L'analyse bayésienne rencontre un succès certain pour l'ajustement de modèles PBPK à des données toxicocinétiques (Bayard and Jelliffe, 2004; Bois *et al.*, 1996a; Bois *et al.*, 1996c; Bortot *et al.*, 2002; Jonsson and Johanson, 2001, 2002; Lunn and Aarons, 1998; Wakefield, 2000).

Cependant, l'ajustement d'un modèle PBPK à des données expérimentales peut se révéler difficile en raison de la complexité du modèle, du nombre important de paramètres à estimer, et de la nature des données toxicocinétiques (généralement peu nombreuses, éparses et entachées d'erreur de mesure). De plus, ces modèles ont tendance à devenir de plus en plus détaillés, donc complexes, notamment dans le cadre d'études du corps entier (Nestorov, 2003), ou encore chez la femme enceinte (Luecke *et al.*, 1994). Pour être fiable et pertinent, l'ajustement de modèles de grande dimension nécessite alors de nombreuses données et informations (Balant and Gex-Fabry, 1990). Pour pallier à ce manque de données, nous proposons dans ce travail de développer des protocoles expérimentaux facilitant la paramétrisation des modèles PBPK. Ces protocoles ont pour objectif d'apporter un complément d'information sur les paramètres des modèles. Pour ce faire, deux axes d'étude sont proposés : la collecte de données sur la distribution spatio-temporelle d'une substance dans l'organisme par des techniques d'imagerie, et l'utilisation de traceurs non-toxiques en complément des expositions à la substance toxique d'intérêt. Ces deux aspects seront ici étudiés de façon préliminaire et exploratoire, afin d'évaluer leur potentialité.

Ce travail se divise en cinq sections. Dans un premier temps, nous introduisons les techniques utilisées, c'est à dire la modélisation PBPK, l'analyse bayésienne, les techniques d'imagerie médicale et l'agrégation de systèmes d'équations différentielles, et détaillons les objectifs spécifiques. Ensuite, les travaux scientifiques publiés dans le cadre de cette thèse sont présentés : l'analyse statistique bayésienne de données toxicocinétiques; la modélisation PBPK de données d'imagerie; et l'utilisation des traceurs non-toxiques pour les études toxicocinétiques. Enfin, une conclusion sur le projet dans sa globalité et ses perspectives est présentée.

### 3. Contexte et objectifs spécifiques

Cette section présente une revue des différentes notions techniques utilisées dans ce travail, et l'état de l'art des connaissances. La modélisation physiologique toxicocinétique est décrite. Nous abordons ensuite la calibration statistique de ces modèles dans un cadre bayésien. Pour finir, nous introduisons l'imagerie par résonance magnétique et les techniques d'agrégation des systèmes d'équations différentielles qui seront utilisées pour le développement des protocoles expérimentaux.

#### 3.1. Les modèles physiologiques toxico/pharmacocinétiques

Le devenir d'une substance dans le corps est communément divisé en quatre phases : l'Absorption, la Distribution, le Métabolisme et l'Excrétion (processus "ADME"). Dans un premier temps, la substance est absorbée, et distribuée dans le corps par un fluide biologique (sang, lymphe, par diffusion...). Ensuite, la substance est soit stockée (par exemple, dans les graisses ou les muscles), soit éliminée (par métabolisme ou excrétion). De par leur description paramétrique de l'organisme, les modèles PBPK s'appuient sur les connaissances biologiques et anatomiques pour décrire ces processus. Plus précisément, le corps est divisé en plusieurs compartiments ayant tous une signification physiologique. Les fluides biologiques, principaux vecteurs des composés dans le corps, créent les interconnexions entre les compartiments.

Les premiers travaux incluant une modélisation mécaniste et physiologique concernent la pharmacocinétique de xénobiotiques et datent des années 1930 (Teorell, 1937). Il a fallu attendre les travaux de Bischoff à la fin des années 1960 (Bischoff and Brown, 1966) pour que les modèles physiologiques deviennent des outils répandus pour l'étude de la cinétique de substances thérapeutiques ou toxiques. Bischoff et Brown (1966) ont proposé une description physiologique très détaillée de la distribution d'une substance au sein d'un organe (ou compartiment), et développé un modèle général de distribution pour les mammifères. Cette approche a ensuite été utilisée pour prédire l'évolution de la concentration d'une substance d'intérêt dans différents organes (Bischoff *et al.*, 1971; Dedrick *et al.*, 1973; Fernandez *et al.*, 1977; Sato *et al.*, 1977). Gerlowski et Jain (1983) présentent une revue des premiers modèles PBPK développés pour des substances thérapeutiques ou toxiques.

Initialement, les modèles PBPK ont été développés dans le but heuristique de comprendre la distribution d'une substance, et de prédire les profils d'évolution de sa concentration dans différents organes ou tissus. Ces modèles permettent, en effet, d'organiser les différentes sources d'information (telles que des données obtenues par des techniques *in vitro*, ou référencées dans la littérature), et d'expliquer les données toxico/pharmacocinétiques observées. L'utilisation des modèles PBPK s'est rapidement développée en toxicologie, notamment pour évaluer les risques encourus après une exposition à un toxique (Andersen, 1995a; Bailer and Dankovic, 1997; Bois, 1999; Clewell, 1995; Jonsson, 2001). Bien que les modèles physiologiques aient été développés initialement pour les substances thérapeutiques, ils ont fait l'objet de nombreuses réserves en pharmacologie (Charnick *et al.*, 1995). Les reproches les plus courants sont la

complexité et la grande dimension de ces modèles (Ings, 1990), ainsi que la difficulté d'identifier et d'estimer leurs paramètres (Veng-Pedersen and Modi, 1992). C'est pourquoi le développement de modèles plus simples (tels que les modèles compartimentaux) a longtemps été préféré dans ce domaine. Récemment, plusieurs auteurs ont démontré l'utilité de ces modèles dans le processus de développement de nouvelles substances thérapeutiques (Blesch *et al.*, 2003; Mather, 2001; Theil *et al.*, 2003).

Dans la suite, nous présentons la construction d'un modèle PBPK : la mise en équation des processus ADME, la spécification du modèle et de ses paramètres. Pour finir, nous discutons de l'intérêt des modèles PBPK pour les études d'évaluation de risques et les études thérapeutiques.

### 3.1.1. Les processus ADME

Dans cette section, nous limitons la description des processus ADME au cas où la substance est distribuée *via* le réseau sanguin. Dans ce cas, un facteur limitant de la distribution au sein d'un organe est la perfusion de cet organe. Les échanges sanguins sont le plus couramment décrits dans les modèles PBPK, mais d'autres échanges (via la lymphe ou par diffusion) peuvent être modélisés.

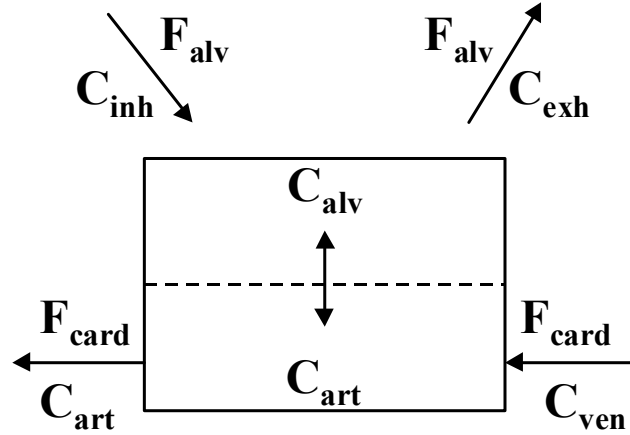
#### A. Absorption

On appelle absorption la pénétration de la substance dans l'organisme. La phase d'absorption dépend du mode d'administration. On peut différencier deux modes d'administration : l'administration intra ou extra-vasculaire. Lorsque l'administration est intra-vasculaire, il n'y a pas de phase d'absorption par l'organisme puisque la substance est directement injectée dans le réseau sanguin. Si l'administration est extra-vasculaire, tout ou une partie de la substance sera absorbée. Les modes d'administration extra-vasculaire principaux sont l'ingestion, l'inhalation, et la pénétration trans-cutanée. Pour quantifier la dose de substance absorbée, il est essentiel de modéliser les principaux déterminants de la relation entre la dose administrée et celle absorbée. Divers facteurs, tels que la solubilité de la substance, la perméabilité des tissus ou encore la durée d'exposition, influencent ce processus. Prenons l'exemple d'une exposition par inhalation. Andersen (1991) a proposé une modélisation des échanges de substance au niveau pulmonaire. Ce modèle divise les poumons en espace alvéolaire et en sang artériel. Il fait l'hypothèse d'un équilibre rapide de la substance entre le sang artériel et l'air alvéolaire, sans métabolisme au niveau pulmonaire, et une capacité limitée de stockage de la substance dans les poumons. La condition d'équilibre est déterminée par le coefficient de partage sang sur air,  $P_{s/a}$ . La Figure 2.1 donne une représentation graphique du modèle, qui se traduit par le système suivant:

$$\begin{aligned} C_{art} &= \frac{P_{s/a} (F_{alv} \times C_{inh} + F_{card} \times C_{ven})}{P_{s/a} \times F_{card} + F_{alv}} \\ C_{exh} &= \frac{C_{art}}{P_{s/a}} \end{aligned} \quad (2.1)$$

$C_{alv}$ ,  $C_{art}$ ,  $C_{ven}$ ,  $C_{inh}$  et  $C_{exh}$  sont respectivement les concentrations dans le sang alvéolaire, artériel et veineux, et dans l'air inhalé et exhalé.  $F_{card}$  et  $F_{alv}$  sont respectivement les flux sanguins cardiaque et alvéolaire.

Pour une même exposition, il peut y avoir plusieurs sites d'administration. En effet, si on souhaite modéliser l'exposition d'un sujet à un gaz se trouvant dans l'air ambiant, il faudra tenir compte évidemment de l'inhalation, mais aussi éventuellement de la pénétration trans-cutanée (Corley *et al.*, 1997).



**Figure 2.1 :** Modélisation des échanges d'une substance au niveau pulmonaire.  $C_{alv}$ ,  $C_{art}$ ,  $C_{ven}$ ,  $C_{inh}$  et  $C_{exh}$  sont respectivement les concentrations dans le sang alvéolaire, artériel et veineux, et dans l'air inhalé et exhalé.  $F_{card}$  et  $F_{alv}$  sont respectivement les flux sanguins cardiaque et alvéolaire. Les flèches schématisent les échanges de matière.

## B. Distribution

La distribution concerne la répartition de la substance dans les différents organes du corps. Une fois absorbée, la substance passe du sang vers les organes selon son affinité pour eux. Ce phénomène est caractérisé par un paramètre appelé coefficient de partage. Cependant, d'autres facteurs déterminent aussi la dynamique de distribution de la substance. Ces facteurs sont en particulier la taille des organes, la perméabilité des membranes des tissus ou encore la perfusion des organes. La Figure 2.2 présente la distribution d'une substance dans un tissu, dans lequel il n'y a pas de phénomène d'élimination ou d'absorption. Notons  $Q_A$ ,  $Q_V$  et  $Q_T$  les quantités dans le sang artériel et veineux, et dans le tissu respectivement.  $V_A$ ,  $V_V$  et  $V_T$  sont respectivement les volumes du sang artériel et veineux, et du tissu.  $Flux$  correspond au flux sanguin. La substance est présentée au tissu avec un débit massique égal à  $Flux \times Q_A / V_A$  et quitte le tissu avec le débit massique  $Flux \times Q_V / V_V$ . La quantité de substance dans le tissu au temps  $t$  est donc donnée par :

$$\frac{\partial Q_T(t)}{\partial t} = Flux \times \left( \frac{Q_A(t)}{V_A} - \frac{Q_V(t)}{V_V} \right) \quad (2.2)$$

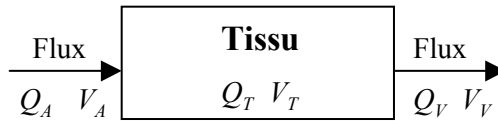
Par ailleurs, la concentration veineuse à la sortie du tissu et la concentration tissulaire sont liées par la relation suivante :



$$C_V(t) = \frac{C_T(t)}{PC_{T/S}} \quad (2.3)$$

où  $PC_{T/S}$  est le coefficient de partage tissu sur sang de la substance étudiée. Ce coefficient est défini comme étant le rapport de la concentration de la substance dans le tissu sur celle dans le sang lorsque ces deux milieux sont à l'équilibre. L'équation (2.2) peut alors se réécrire:

$$\frac{\partial Q_T(t)}{\partial t} = Flux \times \left( \frac{Q_A(t)}{V_A} - \frac{Q_T(t)}{V_T \times PC_{T/S}} \right) \quad (2.4)$$



**Figure 2.2 :** Distribution d'une substance dans un compartiment.  $Q_A$ ,  $Q_V$  et  $Q_T$  sont respectivement les quantités dans le sang artériel et veineux et dans le tissu.  $V_A$ ,  $V_V$  et  $V_T$  sont respectivement les volumes du sang artériel, veineux et tissulaire. Les flèches schématisent la circulation sanguine.

### C. Métabolisme

Le métabolisme est, schématiquement, le processus de transformation chimique de la substance par les enzymes au sein d'un organe. Les produits ainsi formés sont appelés métabolites. Dans de nombreux cas, ce n'est pas le produit parent (*i.e.*, la substance administrée) qui est actif (toxique ou thérapeutique), mais ses métabolites. On parle de métabolisme de premier passage lorsqu'une fraction de substance est métabolisée au moment de l'absorption, avant d'être distribuée dans l'organisme. La quantité d'enzymes présente dans un organe peut devenir un facteur limitant du métabolisme, notamment pour des expositions à de fortes doses. Un tel processus à capacité limitée est souvent modélisé par une expression de type Michaelis-Menten. L'échange de matière au sein du compartiment se modélisera par :

$$\frac{\partial Q_T(t)}{\partial t} = Flux \times \left( \frac{Q_A(t)}{V_A} - \frac{Q_T(t)}{V_T \times PC_{T/S}} \right) - \frac{V_{max} \times Q_T(t)}{K_m \times V_T + Q_T(t)} \quad (2.5)$$

où  $V_{max}$  représente la vitesse maximum de métabolisme, proportionnelle à la concentration totale d'enzymes dans l'organe considéré.  $K_m$  est une mesure de l'affinité entre la substance et les enzymes. Dans les cas les plus simples, tels que pour des expositions à faibles doses d'une substance (Bois *et al.*, 1999b), les réactions métaboliques peuvent être modélisées par des cinétiques de premier ordre, ce qui donne:

$$\frac{\partial Q_T(t)}{\partial t} = Flux \times \left( \frac{Q_A(t)}{V_A} - \frac{Q_T(t)}{V_T \times PC_{T/S}} \right) - k_T \frac{Q_T(t)}{V_T} \quad (2.6)$$

où  $k_T$  est une constante. Généralement, le site principal de métabolisme est le foie. Pour certaines substances, des organes tels que les reins ou les poumons peuvent aussi avoir un rôle métabolique important.

## D. Excrétion

La seconde voie d'élimination est l'excrétion. Ce phénomène correspond à l'élimination de la substance de l'organisme (les métabolites peuvent aussi être excrétés). Les voies d'élimination peuvent être la voie rénale, fécale, l'exhalation, la transpiration ou encore par le lait maternel. Les reins sont en général l'organe principal pour l'excrétion. Comme pour le métabolisme, ce processus est souvent modélisé par une approximation du premier ordre (équation 2.6) ou par des cinétiques de Michaelis-Menten (équation 2.5).

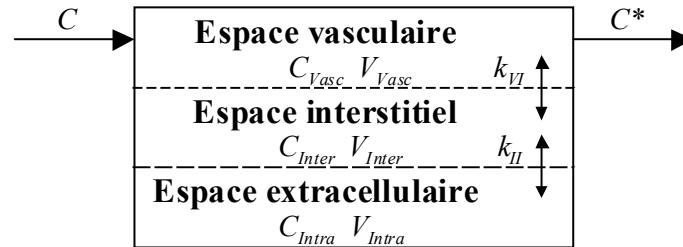
### 3.1.2. Spécification d'un modèle PBPK

Comme dans toute entreprise de modélisation, il est nécessaire de recourir à un grand nombre de simplifications, pour construire un modèle PBPK. La première étape consiste à définir les compartiments nécessaires à la modélisation de la biodistribution de la substance étudiée. Un compromis doit être trouvé entre l'exigence mathématique de décrire le plus finement possible l'organisme et la complexité mathématique du système d'équations différentielles résultant. Il est primordial que les organes ou tissus essentiels à la pharmacocinétique et pharmacodynamique de la substance soient représentés (Nestorov, 2003). Le site d'absorption ou d'administration (les poumons pour une inhalation, la peau pour une exposition dermique, ou encore le sang pour une injection par intraveineuse) et les sites de métabolisme et d'élimination (le foie, les reins et/ou les poumons) sont typiquement décrits. En complément, sont ajoutés d'autres compartiments représentant des groupes généraux de tissus : les tissus bien perfusés (les viscères), les tissus peu perfusés (les muscles, la peau) et les graisses. Certains tissus peuvent être agrégés en raison de flux sanguins et de coefficients de partage similaires (Nestorov *et al.*, 1998).

Le nombre de compartiments dépend aussi du but de la modélisation. Pour une même substance différents modèles PBPK peuvent être construits. Prenons l'exemple du 1,3-butadiène. Kohn et Melnick (1993) ont proposé de modéliser sa cinétique par un modèle à six compartiments différenciant les métabolismes hépatique, viscéral et pulmonaire. Plus récemment, Mezzetti *et al.* (2003) ont proposé un modèle à trois compartiments, dans lequel le compartiment agrégeant les organes bien perfusés est le seul site de métabolisme. Malgré leurs structures différentes, ses modèles ont permis de répondre à des questions toxicologiques précises concernant la même substance.

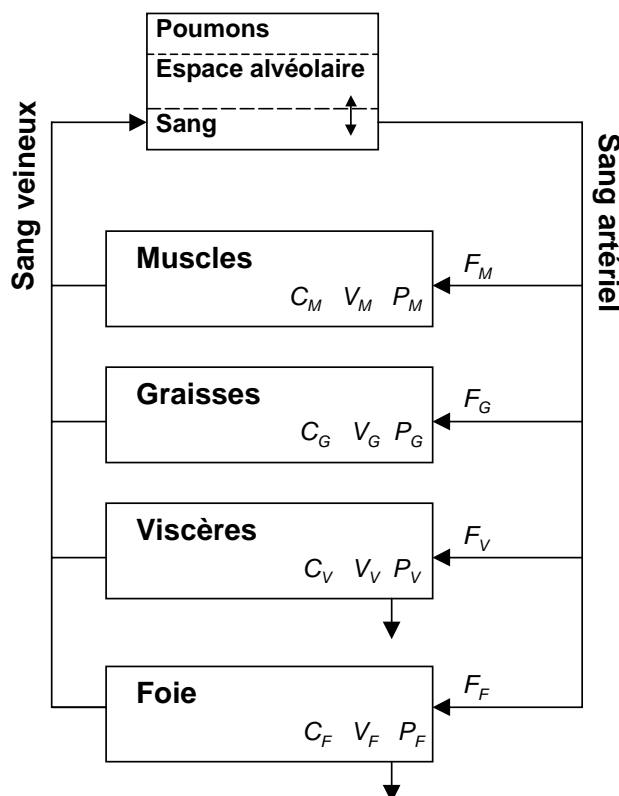
Au sein d'un même compartiment, la substance est supposée se distribuer instantanément et uniformément dans le volume total. Si l'on souhaite décrire des limitations par diffusion (entre espace cellulaire et extra-cellulaire, par exemple), il est possible de subdiviser le compartiment en trois sous-unités (Bischoff and Brown, 1966; Gerlowski and Jain, 1983) : l'espace vasculaire directement perfusé par le sang, l'espace interstitiel qui forme une matrice pour les cellules du tissu, et l'espace cellulaire (Figure 2.3). La substance franchit le mur capillaire et se diffuse dans l'espace interstitiel. Puis elle traverse éventuellement la membrane cellulaire pour rejoindre l'espace intracellulaire. Dans chacun des sous-compartiments, la substance peut être dégradée par une action métabolique ou être captée par une substance endogène. La

quantité de substance peut alors être décrite par un ensemble d'équations modélisant la quantité dans chacun des sous-compartiments.



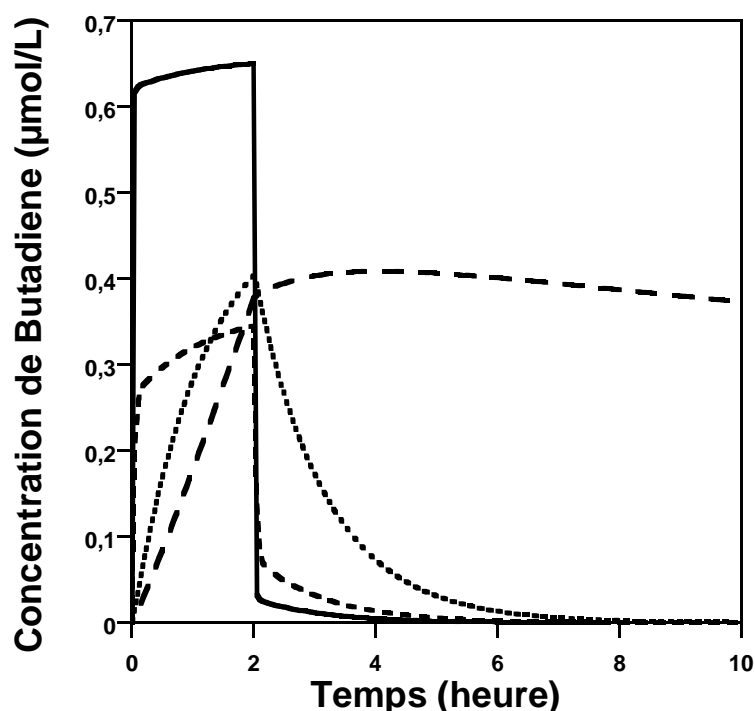
**Figure 2.3 :** Représentation schématique des espaces vasculaire, interstitiel et intracellulaire d'un organe. Chaque sous-compartiment est défini par sa concentration de substance ( $C_{Vasc}$ ,  $C_{Inter}$ ,  $C_{Intra}$ ) et son volume ( $V_{Vasc}$ ,  $V_{Inter}$ ,  $V_{Intra}$ ). Les constantes de transfert entre ces sous-compartiments sont  $k_{VI}$  et  $k_{II}$ .  $C$  et  $C^*$  sont respectivement les concentrations de substance entrant et sortant de l'organe. Les flèches représentent les échanges de matière.

La Figure 2.4 présente un modèle PBPK classique, utilisé notamment pour décrire les cinétiques du 1,3-butadiène chez les mammifères (Kohn and Melnick, 1993). Le 1,3-butadiène est administré par inhalation. Ce type d'administration a nécessité une modélisation des échanges au niveau pulmonaire, notamment pour décrire le passage du 1,3-butadiène de l'air exhalé dans la circulation sanguine. L'excrétion et le métabolisme ont lieu dans les poumons, le foie et les reins. Dans chacun de ces trois compartiments, le métabolisme est décrit par des équations de Michaelis-Menten.



**Figure 2.4 :** *Modèle PBPK pour décrire la cinétique du 1,3-butadiène chez les mammifères, proposé par Kohn and Melnick (1993). Six compartiments sont définis : les poumons (site d'administration, d'élimination et de métabolisme), les viscères (site de métabolisme), les muscles, les graisses, le foie (site de métabolisme) et le sang. Les compartiments sont interconnectés par le réseau sanguin. Chaque compartiment est décrit par son volume ( $V$ ), sa concentration de 1,3-butadiène ( $C$ ), son coefficient de partage ( $P$ ) et son flux sanguin ( $F$ ).*

La Figure 2.5 présente la prédiction de l'évolution temporelle de la concentration de 1,3-butadiène dans quatre compartiments (les poumons, viscères, muscles et graisses) pour une exposition à 5 ppm (parties par million, en volume) de 1,3-butadiène pendant 2 heures. Les valeurs des paramètres ont été obtenues dans la littérature et correspondent à la physiologie d'un homme standard (Kohn and Melnick, 1993). Sur cette figure, on distingue différents types de profil, qui dépendent de la nature de l'organe considéré, c'est à dire si l'organe est un site d'administration, d'élimination, de métabolisme, ou encore d'accumulation.



**Figure 2.5 :** Evolution de la concentration de 1,3-butadiène dans quatre compartiments du modèle PBPK proposé par Kohn and Melnick (1993): les poumons (ligne pleine), les viscères (tirets courts), les muscles (ligne pointillée), et les graisses (tirets longs). Une exposition à 5 ppm de 1,3-butadiène pendant 2 heures a été simulée pour un homme standard.

### 3.1.3. Paramétrisation d'un modèle PBPK

Comme nous l'avons vu dans les sections précédentes, la construction d'un modèle PBPK dépend directement de l'organisme étudié, mais aussi de la substance. En conséquence, l'ensemble des paramètres d'un modèle PBPK inclut trois sous-ensembles distincts (Andersen, 1995b) :

- les paramètres relatifs à l'anatomie du sujet et aux processus physiologiques sous-jacents, tels que les volumes des organes ou les flux sanguins,
- les paramètres biochimiques spécifiques à la substance étudiée et caractéristiques de ses propriétés pharmacocinétiques, tels que les coefficients de partage ou les constantes d'élimination,
- les paramètres relatifs à l'exposition, c'est-à-dire la dose et la durée d'exposition. Ces paramètres de design expérimental n'ont de sens qu'associés à un modèle statistique, ou pour réaliser des simulations.

Chaque compartiment est au moins défini par trois paramètres : son flux sanguin, son volume et le coefficient de partage relatif à la substance. Les modèles PBPK comportent donc de nombreux paramètres.

Toutefois, il est possible d'exprimer certains paramètres physiologiques en fonction de caractéristiques individuelles des sujets étudiés. Typiquement, l'âge, le sexe, la masse corporelle et la taille sont des covariables enregistrées, notamment pour les sujets humains. Ensuite, des algorithmes permettent de lier ces covariables à des paramètres du modèle. Par exemple, Deurenberg *et al.* (1991) proposent de calculer la masse de graisse d'un sujet humain à partir des quatre covariables citées précédemment. De même, les volumes des organes sont généralement mis à l'échelle de la masse du sujet considéré et les flux des organes mis à l'échelle du flux total (Bernillon and Bois, 2000). Dans ce cadre, le volume du compartiment  $i$  ( $V_i$ ) sera donné par la relation suivante :

$$V_i = sc\_V_i \times MC \quad (2.7)$$

où  $sc\_V_i$  est le paramètre d'échelle et  $MC$  la masse corporelle du sujet. Ce type d'équation simplifie la paramétrisation du modèle pour deux raisons principales. La première est que les valeurs de paramètres référencées dans la littérature sont typiquement exprimées en fonction de la masse ou du flux total (Brown *et al.*, 1997; International Commission on Radiological Protection (ICRP), 2003). La seconde concerne la réduction du nombre de paramètres à estimer. En effet, il est possible de contraindre la somme des paramètres d'échelle des flux sanguins et des volumes à 1 (ou tout autre nombre si tous les organes ne sont pas représentés, par exemple), et de ce fait d'éviter l'estimation d'un paramètre. Par exemple, supposons qu'un modèle contienne 3 compartiments ( $i$ ,  $j$  et  $k$ ). Notons  $sc\_V_i$ ,  $sc\_V_j$  et  $sc\_V_k$  les paramètres d'échelle des volumes. Pour réduire le nombre de paramètre à estimer, il suffira de définir  $sc\_V_k$  en fonction de  $sc\_V_i$  et  $sc\_V_j$ , c'est-à-dire :

$$sc\_V_k = 1 - (sc\_V_i + sc\_V_j) \quad (2.8)$$

### 3.1.4. Un outil pour l'évaluation de risque et les études thérapeutiques

De par leur description de l'organisme, les modèles PBPK fournissent un cadre paramétrique nécessaire aux extrapolations et transpositions inhérentes aux études toxico/pharmacocinétiques. Pour étudier les effets associés à l'exposition à une substance, des études expérimentales sont en général nécessaires. Pour des raisons éthiques évidentes, il est impossible d'exposer des volontaires humains à des toxiques. Les expositions de volontaires humains sont uniquement autorisées à la fin du développement des substances thérapeutiques, ou à des valeurs environnementales d'exposition pour des toxiques. Les expérimentations animales deviennent alors utiles pour comprendre et caractériser les mécanismes biologiques sous-jacents de l'action de la substance. Les résultats obtenus doivent ensuite être extrapolés aux conditions réelles d'exposition pour des populations humaines. Pour toutes ces extrapolations ou transpositions, un principe prévaut : la forme analytique de la cinétique de la substance dans le corps est supposée identique quelque soit les conditions expérimentales, les espèces ou individus exposés. Seules les valeurs des paramètres du modèle expliquent les différences observées entre les effets obtenus sur des espèces ou individus différents.

### A. Transpositions entre conditions expérimentales

Les conditions expérimentales utilisées lors d'expositions en laboratoire sont bien souvent éloignées de celles rencontrées dans notre vie quotidienne. L'exposition à une substance peut être continue mais uniquement sur le lieu de travail, ou discontinue lors de la prise à temps fixe d'un médicament, par exemple. La dose, la concentration et le temps d'exposition sont des paramètres du modèle PBPK. Chaque modèle est uniquement valide sur un certain domaine, appelé domaine de validité. Pour transposer à de nouvelles conditions expérimentales, il faut donc s'assurer qu'elles appartiennent bien au domaine de validité. Pour extrapoler à de nouveaux temps ou doses d'exposition, il suffit de modifier la valeur de ces paramètres pour prédire le comportement de la substance selon ces nouvelles conditions expérimentales (Gelman *et al.*, 1996b). Une extrapolation entre faibles et fortes doses reste cependant délicate. Il est nécessaire de vérifier que la modélisation de certains phénomènes, tels que le métabolisme et l'élimination, soit en accord avec les nouvelles conditions expérimentales. Si le métabolisme a été modélisé par un paramètre de premier ordre dans le cadre d'exposition à faibles doses, il faudra vérifier que cette hypothèse est encore valable pour des fortes doses.

En plus de ces extrapolations inter-durée et inter-dose, il est possible de transposer entre voies d'administration. Une fois dans le corps, la substance se distribue selon les mêmes mécanismes indépendamment de la voie d'administration. Pour extrapoler à une autre voie d'administration, il est simplement nécessaire de modéliser la nouvelle phase d'absorption pour quantifier la dose de substance atteignant le fluide biologique la véhiculant dans les différents organes du corps (le réseau sanguin, par exemple).

### B. Transpositions inter-espèce, inter- et intra-individu

Une fois le modèle déterminé à l'aide de données animales, il devient possible de transposer les résultats obtenus chez l'animal à l'humain (ou autre espèce cible). Pour cela, il suffit d'affecter aux paramètres physiologiques du modèle des valeurs concernant la nouvelle espèce (Watanabe and Bois, 1996; Welsch *et al.*, 1995). Comme nous l'avons souligné précédemment, ces valeurs peuvent être trouvées dans la littérature. Cependant, certains paramètres ne sont pas référencés pour toutes les espèces, tels les coefficients de partage ou de diffusion d'un compartiment à l'autre. Pour y remédier, une solution est le recours aux équations allométriques (Singer, 2001; West *et al.*, 1997). Un exemple d'équation allométrique est :

$$P = a \times MC^k \quad (2.9)$$

où  $P$  est la valeur extrapolée du paramètre considéré,  $a$  et  $k$  des paramètres indépendants des espèces et  $MC$  la masse corporelle de l'espèce considérée. Ces équations sont en général paramétrées par ajustement des paramètres  $a$  et  $k$  sur la base de données expérimentales (*in vitro*, par exemple) ou provenant de la littérature. D'autres techniques, de type QSAR, ont été développées pour prédire les coefficients de partage tissus sur sang, par exemple (Poulin and Theil, 2000).

De manière identique, il est possible de transposer des résultats obtenus chez une population standard à des populations à risque, telles que les femmes enceintes ou les enfants. Comme pour les transpositions inter-espèces, il suffit d'assigner aux paramètres du modèle des valeurs tenant compte des caractéristiques particulières de ces populations.

La dernière transposition que nous traitons est l'intra-individuelle. L'anatomie et la physiologie d'un même individu évoluent au cours de sa vie. Là encore, les valeurs des paramètres du modèles peuvent expliquer ces différences. Cette technique de transposition intra-individuelle a permis de prédire l'accumulation de 2,3,7,8-tetrachloro-*p*-dioxine (TCDD) à long terme selon l'évolution de la masse des tissus chez l'homme et la femme (Van der Mollen *et al.*, 1996). De même, plusieurs modèles physiologiques ont été développés pour décrire les changements biologiques survenant chez la femme lors d'une grossesse (Luecke *et al.*, 1994).



## 3.2. Calibration et analyse bayésienne

Dans cette section, nous nous intéressons à l'ajustement de modèles PBPK à des données expérimentales. Comme nous l'avons vu précédemment, ces modèles sont définis à l'aide de nombreux paramètres physiologiques et anatomiques. Il est connu que les paramètres biologiques (tels que les volumes des organes, les flux sanguins, les coefficients de partage, etc.) peuvent être très hétérogènes au sein même d'une population. Pour traduire cette hétérogénéité, une solution consiste à définir une distribution de probabilité pour chaque paramètre, à partir de valeurs référencées dans la littérature ou obtenues expérimentalement. Cette distribution doit refléter la connaissance sur chaque paramètre. Le cadre d'inférence statistique permettant de tenir compte naturellement de ces distributions dans le processus d'estimation des paramètres est le cadre bayésien (Bernardo and Smith, 1994; Driesbeke *et al.*, 2002; Gelman *et al.*, 1995). Dans un premier temps, nous présentons le principe d'inférence statistique bayésienne. Ensuite les différents aspects techniques sont abordés, tels que les modèles statistiques utilisés et les algorithmes de simulations permettant l'estimation des paramètres par l'ajustement du modèle aux données (appelé calibration dans un contexte bayésien). Dans une partie suivante, un processus complet d'analyse bayésienne de données toxicocinétiques sera présenté plus longuement.

### 3.2.1. L'inférence bayésienne

Dans un contexte bayésien, toutes les quantités inconnues sont considérées comme des variables aléatoires, en particulier les paramètres du modèle (ensemble noté  $\theta$ ). Avant d'analyser les données expérimentales, une distribution de probabilité *a priori*,  $P_\theta$ , est proposée pour chaque paramètre inconnu. Cette distribution peut être interprétée en terme de degré de croyance sur les valeurs possibles de  $\theta$ , et reflète la connaissance actuelle sur  $\theta$ .

Les distributions *a priori* sont ensuite mises à jour avec les nouvelles données expérimentales, pour conduire à une distribution de probabilité *a posteriori*. Cette mise à jour est réalisée selon le théorème de Bayes qui stipule que la distribution de probabilité *a posteriori*  $P(\theta|y)$  est proportionnelle au produit de toutes les distributions *a priori*,  $P_\theta$ , et de la vraisemblance des données,  $P(y|\theta)$  :

$$P(\theta|y) = \frac{P_\theta \times P(y|\theta)}{P(y)} \quad (2.10)$$

Le dénominateur  $P(y)$ , indépendant de  $\theta$ , peut être interprétée comme une constante nécessaire pour normaliser  $P(\theta|y)$ . La distribution de probabilité *a posteriori* obtenue représente la connaissance sur les paramètres, incluant l'information contenue dans les données. Diverses statistiques sur les paramètres (par exemple, la moyenne, l'écart type, les quantiles) peuvent ensuite être obtenues des distributions *a posteriori*.

Dans les cas les plus simples, tels que les modèles linéaires, l'expression analytique de la loi *a posteriori*  $P(\theta|y)$  peut être déterminée. Cependant, ces cas sont relativement rares en pratique, et il est

souvent difficile d'obtenir une expression analytique pour  $P(\theta|y)$  en raison des termes d'intégration multidimensionnelle contenus dans les expressions de ces lois. Les méthodes Monte Carlo par Chaînes de Markov (MCMC) ont permis de lever un certain nombre de ces difficultés. Ces méthodes fournissent des échantillons des valeurs de paramètres tirés de leur distribution *a posteriori* même si une expression analytique ne peut être obtenue.

### 3.2.2. Modèles statistiques

Etant donnée l'équation 2.10, la spécification d'un modèle statistique bayésien se décompose en deux éléments: le choix des distributions *a priori* des paramètres, et la définition de la vraisemblance.

#### A. Distributions *a priori*

Une distribution *a priori* est définie pour chacun des paramètres du modèle. Cette distribution doit traduire les connaissances sur le paramètre avant l'observation des données à analyser. Son indépendance par rapport aux données est obligatoire afin d'éviter d'utiliser deux fois la même information. La connaissance *a priori* inclut, par exemple, des données développées *in vitro*, des informations obtenues dans la littérature ou par l'expertise. Suivant la disponibilité de ces informations, on distingue deux types de distributions *a priori* : les distributions informatives et non-informatives.

Comme leur nom l'indique, les distributions non-informatives (ou vagues) sont utilisées quand peu d'information est disponible sur la valeur des paramètres. Ces distributions ont donc une forme très aplatie, et sont souvent caractérisées par des lois uniformes. Typiquement, les distributions non-informatives sont définies pour les paramètres spécifiques de la substance étudiée, tels que les constantes de diffusion d'un compartiment à l'autre. A l'inverse, des distributions informatives ou précises peuvent être utilisées pour les paramètres dont les valeurs possibles sont bien connues, tels que les flux sanguins ou les volumes d'organe par exemple. En effet, de nombreux ouvrages rassemblant des caractéristiques anatomiques et physiologiques de différentes espèces sont disponibles (Arms and Travis, 1988; Brown *et al.*, 1997; International Commission on Radiological Protection (ICRP), 2003). En plus des valeurs moyennes, quelques études fournissent aussi la dispersion de la valeur du paramètre d'intérêt dans la population. Avec ces deux informations (moyenne et dispersion), il est possible de construire une distribution *a priori* relativement précise. Pour des valeurs physiologiques et anatomiques, ces distributions prennent souvent la forme de distributions normales ou log-normales.

#### B. Vraisemblance

La forme de la vraisemblance,  $P(y|\theta)$ , dépend d'une part du modèle déterministe (le modèle pharmacocinétique dans notre cas), mais aussi du modèle d'erreur sur les données. En général, le modèle est calibré à l'aide de mesures expérimentales entachées d'erreur qu'il est nécessaire de prendre en compte. Dans le cadre d'études toxico/pharmacocinétiques, ces mesures correspondent souvent à des concentrations, par exemple dans l'air exhalé ou dans les urines. Il est courant de recourir au modèle log-normal pour décrire

l'erreur de telles mesures. Pour définir la variance du terme d'erreur, deux approches sont possibles : soit la variance est constante (Jonsson and Johanson, 2002; Mezzetti *et al.*, 2003), soit elle est dépendante de la concentration mesurée (Bernillon, 2000; Wakefield and Racine-Poon, 1995). Une variance d'erreur variable est souvent définie par un terme absolu, et un terme relatif. Le choix du modèle de variance dépend essentiellement de l'appareil de mesure utilisé. Par exemple, pour les techniques de chimie analytique (telles que la chromatographie gazeuse), les expérimentateurs ont souvent remarqué que la précision des mesures augmentent avec la valeur de la concentration. Pour finir, la variance est soit fixée (à l'aide d'expériences précédentes, par exemple), soit estimée lors de la calibration du modèle.

### 3.2.3. Mise en œuvre de la calibration bayésienne

Dans le cadre d'une calibration bayésienne, nous souhaitons obtenir la distribution *a posteriori* des paramètres  $P(\theta|y)$ , sans avoir à déterminer sa forme analytique. Pour ce faire, les méthodes de simulations Monte Carlo par Chaînes de Markov (MCMC) sont utilisées. Ces méthodes permettent de générer des échantillons d'une fonction dont la forme analytique n'est pas déterminée (Gelman and Rubin, 1996; Gilks *et al.*, 1996; Robert, 1996).

#### A. Méthodes Monte Carlo par chaînes de Markov

Le but des méthodes MCMC est d'obtenir des tirages aléatoires d'une loi de probabilité  $f(x)$  dont l'expression est connue à une constante  $c$  près :

$$f(x) = c \times \varphi(x) \quad (2.11)$$

Ce problème est typiquement présent dans les calculs d'inférence bayésienne, où l'expression de la loi *a posteriori* des paramètres est en général inconnue. Dans l'équation 2.10, le dénominateur  $P(y)$  peut être considéré comme une constante puisque ne dépendant pas de  $\theta$ . En revanche, l'expression du numérateur, fonction de  $\theta$ , est connue puisqu'il s'agit du produit entre la loi *a priori* de  $\theta$ ,  $P_\theta$ , et de la vraisemblance du modèle,  $P(y|\theta)$ .

Un algorithme MCMC est une méthode itérative générant une chaîne de Markov ergodique. Une propriété intéressante des chaînes de Markov ergodiques réside dans le fait qu'elles convergent vers une loi stationnaire, indépendamment des conditions initiales. L'algorithme MCMC est construit de telle sorte que la loi stationnaire corresponde à notre loi d'intérêt, ici  $P(\theta|y)$ . Par définition, les échantillons sont tirés de manière séquentielle, et la distribution dans laquelle la nouvelle valeur est tirée dépend de la valeur obtenue à l'itération précédente. Il existe de nombreuses manières de construire de telles chaînes. Dans la suite, nous nous intéressons à l'algorithme Metropolis-Hastings utilisé dans ce travail, et au problème du contrôle de la convergence des chaînes. Pour finir, une présentation rapide du logiciel utilisé, *MCSim* (Bois and Maszle, 1997b), pour mener nos analyses bayésiennes est proposée.

## B. L'algorithme de Metropolis-Hastings

L'algorithme de Metropolis-Hastings a été développé par Metropolis *et al.* (1953), puis généralisé par Hastings (1970). L'algorithme crée une séquence de vecteurs aléatoires  $(\theta^{(1)}, \theta^{(2)}, \dots)$  dont la distribution conjointe converge vers la loi à échantillonner, ici  $P(\theta|y)$ . Il se décompose selon les étapes suivantes :

1. Une valeur initiale  $\theta^{(0)}$  est tirée dans une distribution initiale  $P_0(\theta)$ ,
2. Pour  $t = 1, 2, \dots$ 
  - (a) un candidat  $\theta'$  est tiré aléatoirement dans une distribution de probabilité  $J_t(\theta|\theta^{t-1})$
  - (b) On calcule le ratio pondéré suivant:

$$r = \frac{P(\theta'|y)/J_t(\theta'|\theta^{t-1})}{P(\theta^{t-1}|y)/J_t(\theta^{t-1}|\theta')} \quad (2.12)$$

- (c) Si  $r \geq 1$ , alors  $\theta^t = \theta'$ , sinon:
  - on tire un nombre aléatoire  $u$  dans une distribution uniforme sur  $[0; 1]$
  - si  $u \leq r$ , on pose  $\theta^t = \theta'$ , sinon  $\theta^t = \theta^{t-1}$

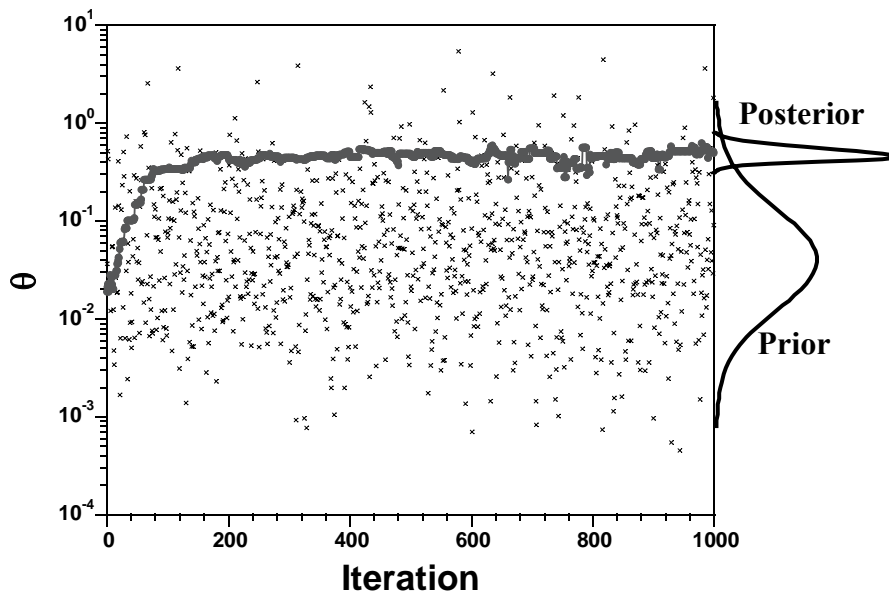
La Figure 2.6 présente l'échantillonnage d'un paramètre par une l'algorithme de Metropolis-Hastings (correspondant à sa distribution *a posteriori*) ainsi que sa distribution *a priori*.

$J_t(\theta|\theta^{t-1})$  est appelée "distribution de saut" car le passage à une nouvelle valeur de tentative, à partir de la dernière valeur retenue, peut être interprété comme un saut dans l'espace des valeurs possibles des paramètres. Sous conditions de régularité peu restrictives, la convergence de la chaîne vers la distribution cible est assurée indépendamment du choix de la fonction de saut. Cependant, même si la convergence théorique n'est pas mise en cause, en pratique, le choix des distributions de saut est un moment décisif dans la mise en place de la méthode. Gelman *et al.* (1995) ont défini certaines propriétés pour construire une "bonne" distribution de saut:

- échantillonner  $\theta$  à partir de  $J_t(\theta|\theta)$  doit être facile,
- le ratio  $r$  doit être facilement calculable,
- chaque saut doit conduire à une distance suffisamment grande dans l'espace des paramètres,
- les sauts ne doivent pas être rejetés trop fréquemment.

Des distributions trop dispersées, qui génèrent des candidats souvent refusés, peuvent laisser la chaîne longtemps "bloquée" sur une valeur et demander un nombre élevé d'itérations. Inversement, avec des fonctions de saut peu dispersées, la chaîne peut bouger trop lentement. L'effet est le même : trop d'itérations sont nécessaires pour atteindre la convergence.

D'autres algorithmes peuvent être utilisés pour générer une chaîne MCMC, tels que l'échantillonnage de Gibbs (Geman and Geman, 1984). Ce dernier requiert la connaissance de l'expression analytique des lois de distributions conditionnelles de chaque variable du modèle en fonction des autres variables et des données.



**Figure 2.6 :** Comparaison de la distribution *a priori* d'un paramètre  $\theta$  et de sa distribution *a posteriori*, échantillonnée par une méthode de simulation Monte Carlo par chaîne de Markov. L'algorithme utilisé est celui de Metropolis-Hastings.

### C. Contrôle de la convergence

Les méthodes de simulations MCMC permettent de générer des échantillons de la distribution *a posteriori* des paramètres du modèle. Ces méthodes de simulations étant basées sur des résultats théoriques asymptotiques, il convient de s'assurer que la convergence est atteinte à l'arrêt des simulations. En pratique, deux approches sont utilisées lors de la calibration d'un modèle : soit une seule chaîne MCMC est simulée pour un nombre d'itérations pour lequel la convergence est supposée atteinte, soit plusieurs chaînes MCMC sont simulées à partir de points initiaux différents. Dans ce travail, la seconde approche a été choisie. Pour s'assurer de la convergence de plusieurs chaînes, différents algorithmes sont disponibles (Cowles and Carlin, 1996). Ici, seule la méthode proposée par Gelman et Rubin (1992) a été utilisée.

Cette méthode se base sur la comparaison de la variance inter-chaînes (entre les valeurs des différentes chaînes) et de la variance intra-chaîne (entre valeurs d'une même chaîne), par l'intermédiaire d'un critère appelé facteur de réduction potentiel de variance,  $\hat{R}$ . La comparaison s'effectue pour chacun des paramètres indépendamment. Les chaînes sont arrivées à convergence lorsque  $\sqrt{\hat{R}}$  est égal à 1. Cependant, la valeur 1 n'est atteinte qu'asymptotiquement. En pratique, on considère que la convergence est probablement atteinte lorsque  $\sqrt{\hat{R}}$  est inférieur à 1,2.

## D. Logiciels

Divers logiciels sont actuellement disponibles pour effectuer des analyses bayésiennes. Ces programmes permettent, notamment, de mener les calculs d'inférence bayésienne sur des modèles complexes, avec divers algorithmes, tels que l'algorithme de Metropolis-Hastings ou de Gibbs. Un des logiciels les plus connus et les plus utilisés est *BUGS* (pour *Bayesian Inference Using Gibbs Sampler*), ou *WinBUGS*, sa version sous Windows (Spiegelhalter *et al.*, 1996). De nombreux exemples d'application de *WinBUGS* ont été publiés (une liste non exhaustive est disponible sur le site internet du projet <http://www.mrc-bsu.cam.ac.uk/bugs/>).

Dans le cadre de ce travail, nous avons utilisé le logiciel *MCSim* (pour *Monte Carlo Simulation*) développé par Bois et Maszle (1997b; 1997c). Ce logiciel a été utilisé principalement pour analyser des données toxico/pharmacocinétiques, et plus particulièrement dans un contexte de population (Bois, 2000; Bois *et al.*, 1999b; Brochot *et al.*, 2004; Jonsson and Johanson, 2003). En pratique, la calibration d'un modèle pharmacocinétique à des données expérimentales requiert deux étapes. Dans un premier temps, il faut décrire formellement le système d'équations différentielles à résoudre, ainsi que les relations entre les différentes quantités du modèle. Ensuite, il est nécessaire de spécifier les distributions *a priori* des paramètres et la vraisemblance des données. S'ils existent, les niveaux de hiérarchie doivent être définis lors de cette seconde étape.

### 3.3. L'imagerie par résonance magnétique

L'imagerie par résonance magnétique nucléaire (IRM) permet, dans le secteur médical, d'obtenir de façon non-invasive des images de haute qualité de l'intérieur du corps. L'IRM est basée sur le principe de la résonance magnétique nucléaire (RMN), et utilise le fait que certains noyaux atomiques possèdent un moment magnétique. La RMN consiste à détecter les changements de moment magnétique d'un échantillon (biologique, par exemple) sous l'action d'un champ magnétique extérieur et d'une onde électromagnétique appropriée. Ces variations du moment magnétique sont à l'origine d'un signal, le signal RMN, qui correctement interprété fournit des informations sur l'échantillon étudié. Pour obtenir une image de cet échantillon, il faut être en mesure de relier les informations obtenues en analysant le signal RMN émis et la position correspondante dans l'échantillon. En utilisant un champ magnétique variant d'un point à l'autre de l'échantillon d'une façon prédéterminée, on crée une dépendance entre la fréquence et la phase du signal RMN d'une part, et la position dans cet échantillon d'autre part. En décodant le signal, on pourra ainsi associer à chaque point un certain nombre de paramètres caractéristiques de l'échantillon (densité de protons, temps de relaxation, ...) et obtenir une image.

Dans la suite, nous présentons (de manière simplifiée) les différentes étapes nécessaires à la construction d'une image RMN. Au préalable, quelques notions de physique seront rappelées. Une présentation extrêmement détaillée de ces techniques peut être trouvée dans (Kastler *et al.*, 2003; Le Bihan, 1984). En pratique, on s'intéresse surtout au noyau d'hydrogène en raison de sa très grande abondance dans les tissus biologiques et, dans le cadre de cette présentation, nous nous limitons à l'étude de ce noyau. Pour finir, un bref aperçu de plusieurs applications de l'IRM dans la recherche médicale est présenté.

#### 3.3.1. Rappels de physique

Dans cette section, quelques rappels de physique sont présentés. Les notions introduites seront utiles pour comprendre l'origine du signal RMN, et donc la construction d'une image de résonance magnétique.

##### A. Champ et moment magnétiques

On dit qu'un champ magnétique règne dans une région donnée de l'espace si la matière placée dans cet espace peut être soumise à une force d'origine magnétique. Un champ magnétique  $\vec{B}$  est un vecteur caractérisé par son intensité, sa direction et son sens. Prenons par exemple le champ magnétique terrestre. Son intensité varie dans l'espace (à Paris, elle est environ égale à 0,5 Gauss), sa direction est grossièrement celle d'un méridien, et son sens est dirigé du sud vers le nord.

Un moment magnétique  $\vec{M}$  placé dans un champ magnétique est soumis à des forces qui tendent à l'aligner avec le champ magnétique. C'est ce qui se passe pour l'aiguille aimantée de la boussole dans le champ magnétique terrestre.

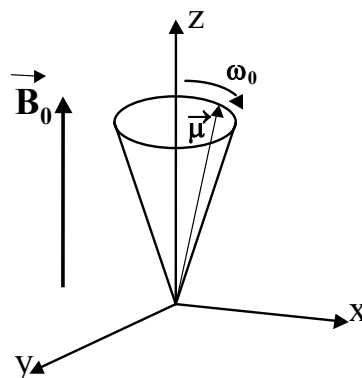
## B. Le principe de résonance

La résonance se définit comme l'existence d'une fréquence pour laquelle un phénomène, tel que la réponse d'un système, soumis à une excitation passe par un maximum. Un système physique (magnétique notamment) est dit résonnant lorsqu'il est susceptible de modifier son état d'équilibre et d'emmagasiner de l'énergie sous l'influence d'une sollicitation externe à une fréquence particulière, appelée fréquence de résonance du système. Après cessation de cette sollicitation, le système revient à sa position d'équilibre, selon des mécanismes oscillants à la fréquence de résonance.

## C. Propriétés magnétiques de la matière

Chaque atome est constitué d'un noyau, lui-même constitué de nucléons (neutrons et protons) et d'électrons gravitant autour de ce noyau. Chaque noyau a une masse et une charge. Certains sont dotés d'un moment cinétique qui caractérise une rotation privilégiée selon un axe et résultante de la répartition des particules du noyau. Le noyau peut donc être assimilé à un aimant microscopique dont la direction du champ est alignée selon son axe de rotation. Les moments cinétiques élémentaires (particules, atomes, molécules) possèdent des propriétés particulières. Le module de ce vecteur est déterminé par un nombre quantique, appelé "spin". Seuls les noyaux ayant un spin non nul (constitués d'un nombre impair de protons et de neutrons) possèdent un moment magnétique dépendant de leurs charges, et sont susceptibles de résonner. Par abus de langage, ces noyaux sont appelés spins.

Chaque composant du noyau (neutrons et protons) possède un moment magnétique intrinsèque. Soumis à un champ magnétique extérieur, chaque spin s'oriente selon l'axe du champ. Une particule qui tourne sur elle-même, en même temps qu'elle s'aligne dans l'axe du champ extérieur, subit un mouvement de rotation autour de l'axe de ce champ : c'est la précession (Figure 2.7). La vitesse de précession d'un spin dépend de la nature de l'atome auquel il appartient et est proportionnelle à l'intensité du champ extérieur.



**Figure 2.7 :** *Mouvement de précession d'un moment magnétique  $\vec{\mu}$  autour d'un champ magnétique  $\vec{B}_0$  avec une vitesse angulaire  $\omega_0$ .*



### 3.3.2. Le signal de résonance magnétique nucléaire

Un signal RMN peut être obtenu en faisant varier les moments cinétiques de la matière. Ici, nous présentons les étapes nécessaires à la création du signal (excitation, relaxation), et les caractéristiques principales du signal résultant.

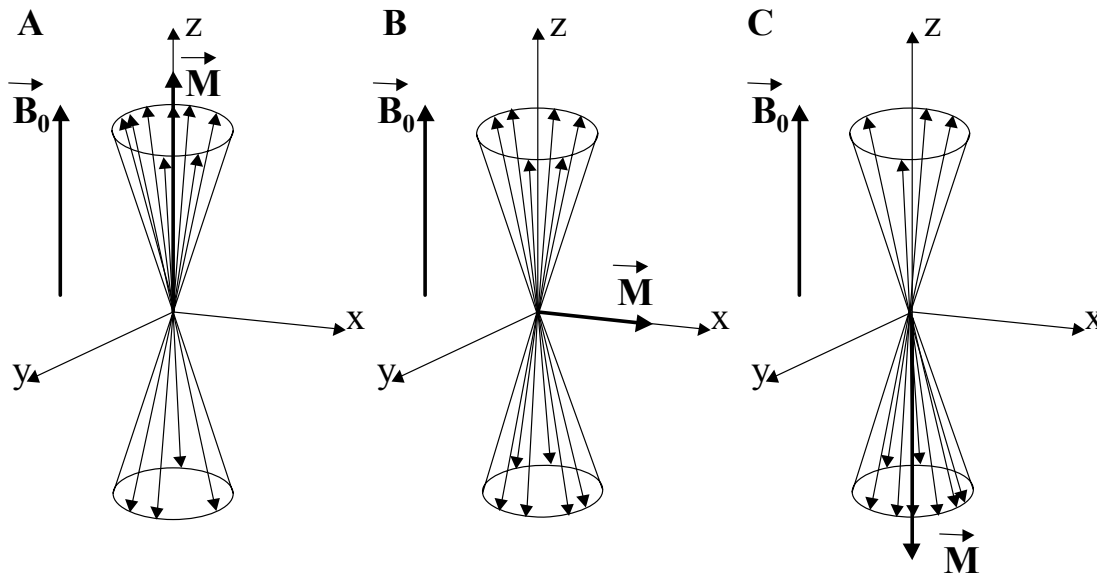
#### A. Action d'un champ magnétique sur une population d'atomes

Le noyau d'hydrogène possède deux niveaux d'énergie :  $E_1$  aussi appelé « Spin  $+1/2$  » et  $E_2$  appelé « Spin  $-1/2$  ». Lorsqu'un noyau d'hydrogène est placé dans un champ magnétique fixe  $\vec{B}_0$ , ces deux niveaux d'énergie sont séparés. Les noyaux d'hydrogène, dirigés dans le sens du champ ou dans le sens contraire, se positionnent sur  $E_1$  ou  $E_2$ . La position parallèle au champ étant préférentielle, les noyaux sont plus nombreux sur le niveau  $E_1$  (niveau le plus stable) que sur  $E_2$ . Notons  $n_1$  et  $n_2$  les nombres de noyaux sur  $E_1$  et  $E_2$  respectivement. Le moment magnétique global  $\vec{M}$  (ou  $M_z$  sa projection sur l'axe du champ) est la somme (en tenant compte du sens) des  $n_1$  moments pointant dans le sens de  $\vec{B}_0$  et des  $n_2$  moments pointant dans le sens contraire. A l'état d'équilibre,  $n_1$  est supérieur à  $n_2$ .  $M_z$  est donc positif et dirigé dans le sens de  $\vec{B}_0$  (Figure 2.8.A). L'aimantation résultante est proportionnelle au nombre de noyaux d'hydrogène.

L'application d'une impulsion électromagnétique à la fréquence de résonance de l'hydrogène va perturber l'aimantation de l'échantillon. Deux cas se distinguent :

- La magnétisation subit une rotation de  $90^\circ$ . On a alors réalisé une « impulsion  $\pi/2$  ». Le vecteur  $\vec{M}$  a totalement basculé dans le plan transversal, c'est-à-dire que la magnétisation longitudinale a disparue ( $M_z = 0$ ). L'amplitude en signal est alors maximale. Une impulsion  $\pi/2$  permet donc d'égaliser les populations des deux niveaux d'énergie ( $n_1 = n_2$ ). On parle alors de « saturation » du signal dans le plan de mesure (Figure 2.8.B).
- La magnétisation subit une rotation de  $180^\circ$ , suite à une « impulsion  $\pi$  ». Le vecteur  $\vec{M}$  s'est alors retourné sur l'axe  $\vec{B}_0$ , et pointe dans la direction opposée à la direction initiale. La composante longitudinale a donc changé de signe, et la composante transversale est nulle. Une impulsion  $\pi$  inverse les populations de  $E_1$  et  $E_2$  (Figure 2.8.C).

La Figure 2.8 récapitule sous forme graphique les différents états cités précédemment : l'état initial, la saturation et l'inversion.



**Figure 2.8 :** *Mouvement de précession d'un ensemble de moments magnétiques nucléaires autour d'un champ magnétique  $\vec{B}_0$ . A l'état initial, l'aimantation résultante  $\vec{M}$  est dans le même sens que  $\vec{B}_0$  en raison d'un nombre de spins orientés vers  $\vec{B}_0$  plus importants (Figure A). Après une impulsion  $\pi/2$ ,  $\vec{M}$  a subi une rotation de  $90^\circ$  (Figure B). Après une impulsion  $\pi$ ,  $\vec{M}$  a subi une rotation de  $180^\circ$  (Figure C).*

## B. Le signal RMN et ses caractéristiques

Après une impulsion électrique, le vecteur  $\vec{M}$  se trouve dans une position ne correspondant pas à son état d'équilibre. Il ne va pas rester dans cette position, et tend à revenir à sa position d'équilibre. Ses composantes transversale et longitudinale ( $M_z$ ) tendent à retrouver leurs valeurs initiales, c'est-à-dire :

- une composante transversale nulle,
- une composante longitudinale  $M_z$  maximale.

Le retour à l'état initial est régi par un processus, appelé la relaxation. Le mouvement complexe effectué par la magnétisation lors du retour à l'équilibre va induire un signal sinusoïdal amorti : c'est le signal RMN.

Prenons l'exemple d'un signal émis après une impulsion  $\pi/2$ . Après cette impulsion, l'aimantation  $\vec{M}$  revient à sa position d'équilibre en un mouvement complexe. Ce dernier peut être décomposé en une composante longitudinale qui va croissant vers sa position d'équilibre ( $\vec{M}_0$ ), et en une composante transversale qui va décroître jusqu'à zéro. Ces mouvements génèrent un signal RMN (Figure 2.9), qui est décrit par son amplitude à l'origine ( $M_0$ ), sa fréquence, et sa constante d'amortissement.

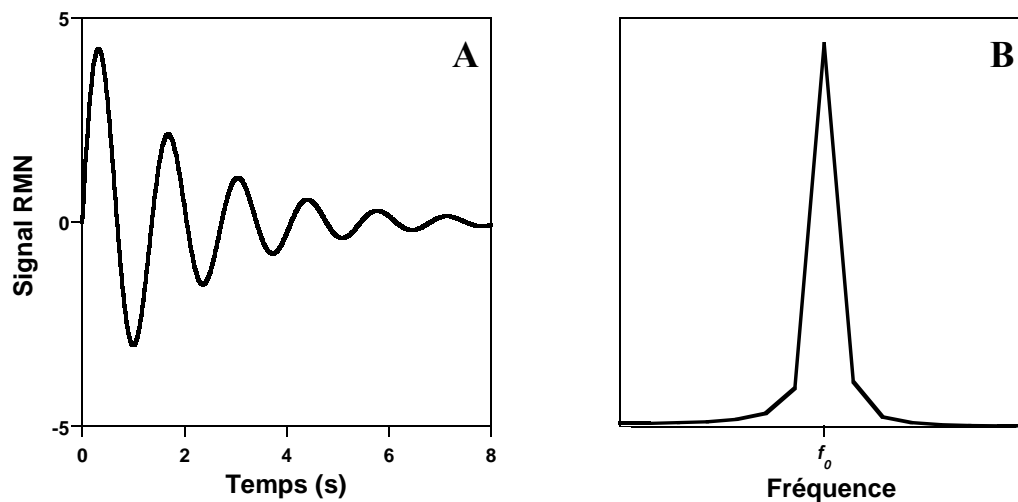
Le signal étant une oscillation amortie, on peut le délimiter à l'intérieur d'une enveloppe, formée par l'évolution de ses limites supérieure et inférieure au cours du temps. L'amplitude du signal est quantifiée par celle de l'enveloppe immédiatement après une impulsion  $\pi/2$ , et est proportionnelle au module de

l'aimantation transversale. Cette mesure permet de représenter la quantité relative de noyaux d'hydrogène par unité de volume.

La fréquence du signal représente le nombre d'oscillations ou cycles par unité de temps. Elle est égale à la fréquence de précession et permet de mesurer l'intensité du champ magnétique. La fréquence du signal RMN sert aussi à localiser ce dernier.

La composante transversale de la précession induit un courant dans une bobine de réception. La durée du signal est donc déterminée par la persistance de l'aimantation transversale qui, en théorie, est infinie. Cependant, en raison des interactions ayant lieu entre les molécules de la matière étudiée, la durée du signal est finie. En pratique, l'intensité du courant induit décroît rapidement au-dessous de la précision de mesure du système de détection. Une constante de temps, représentative de la vitesse de décroissance de l'aimantation transversale, est une caractéristique du tissu : lorsque le champ magnétique est parfaitement uniforme sur l'ensemble de l'échantillon, il s'agit du temps de relaxation transversale  $T_2$ . Certains facteurs peuvent provoquer une décroissance plus rapide du signal, en particulier, une non-uniformité du champ magnétique. En effet, une variation spatiale de l'intensité du champ magnétique se traduit par une variation en fréquence de précession. Les moments magnétiques des différents éléments de volume de l'échantillon ne sont parallèles qu'au début de la précession libre puis, en raison de vitesses de précession différentes, se dispersent dans le plan de mesure. La résultante mesurée par la bobine pour l'ensemble de l'échantillon décroît alors plus vite que chacun des vecteurs. La décroissance du signal est alors caractérisée par le paramètre  $T_2^*$ , plus court que le paramètre  $T_2$  de l'échantillon.

Ce signal RMN peut être représenté non plus en fonction du temps, mais en fonction de sa fréquence. L'opération mathématique qui permet ce changement de variable est la transformée de Fourier. Le signal RMN devient alors une raie de résonance ou un spectre (Figure 2.9).

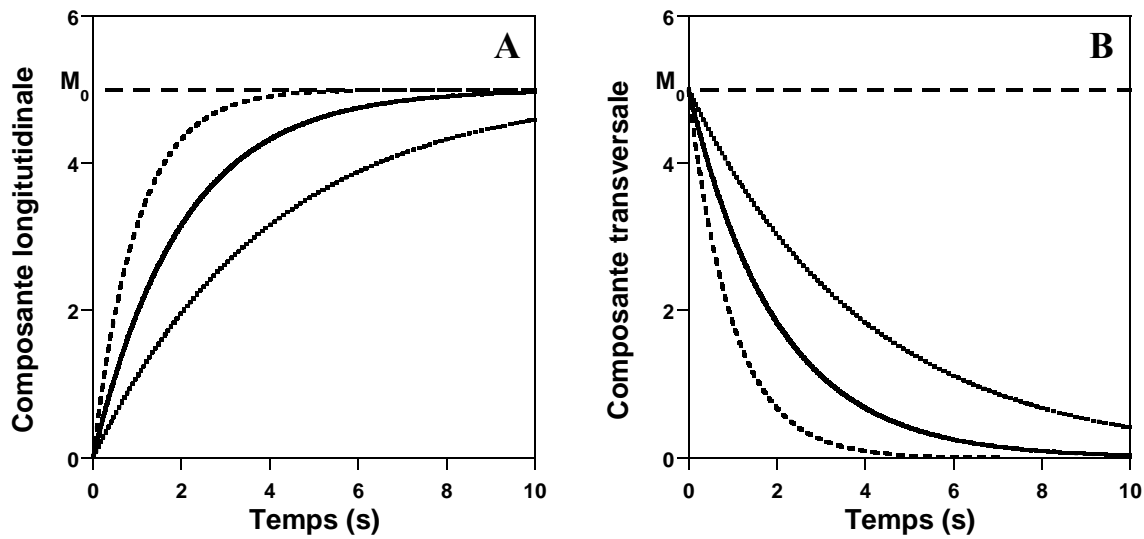


**Figure 2.9 :** *Un signal RMN (A) et son spectre correspondant (B) après une impulsion  $\pi/2$ . Le signal a été obtenu avec les valeurs de paramètre suivantes: amplitude à l'origine ( $M_0$ ) de 5, la vitesse angulaire de 264 rad/s, et la constante d'amortissement ( $T_2^*$ ) de 2 secondes.*

### C. Le phénomène de relaxation

L'importance de la relaxation, c'est-à-dire la rapidité de retour à la valeur initiale, est déterminée par des paramètres temporels : les temps de relaxation  $T_1$  et  $T_2$ .  $T_1$  caractérise la relaxation longitudinale, et  $T_2$  la relaxation transversale. La Figure 2.10 présente la valeur des composantes longitudinale et transversale lors de leur retour à l'équilibre (après une impulsion  $\pi/2$ ) pour différentes valeurs de  $T_1$  et  $T_2$ . Plus  $T_1$  et  $T_2$  sont grands, plus le retour à la position d'équilibre de leurs composantes sera long.

Ces temps de relaxation vont varier pour un tissu donné selon l'organisation des noyaux d'hydrogène. Il est alors possible de s'appuyer sur ces variations pour détecter, au sein d'un même tissu, des modifications liées à la présence d'une lésion.



**Figure 2.10 :** *Courbes des relaxations longitudinale (A) et transversale (B) après une impulsion  $\pi/2$ . Plusieurs valeurs de  $T_1$  et  $T_2$  ont été utilisées pour générer ces courbes. Les lignes pleines correspondent à une valeur de  $T_1$  et  $T_2$  de 2 secondes, les tirets à une valeur de 1 secondes, et la ligne en pointillés à une valeur de 4 secondes. La ligne horizontale marque la valeur de  $M_0$  (ici, 5).*

### 3.3.3. Obtention de l'image de résonance magnétique nucléaire

Nous avons vu comment créer et recueillir un signal RMN provenant de l'ensemble de l'échantillon étudié. Afin de construire une image de cet échantillon tenant compte de son hétérogénéité (par exemple, lorsque différents tissus composent cet échantillon), il est nécessaire de différencier et localiser les signaux émis en des points différents. Pour ce faire, le volume de l'échantillon est découpé en éléments, appelés pixels pour une image en deux dimensions et voxels en trois dimensions. Chaque voxel (ou pixel) est ensuite reproduit à sa place sur l'image RMN par un point dont la brillance correspond à l'intensité du signal RMN émis en ce voxel. L'image ainsi obtenue représente la distribution du signal RMN dans l'échantillon étudié. Pour construire une telle image, il est nécessaire de coder la position de chaque élément dans l'espace, et pour cela de faire varier localement le champ magnétique extérieur  $\vec{B}_0$ .

#### A. Codage de la position du voxel

Le problème qui se pose est la localisation du signal RMN. Pour différencier les points de l'espace, une solution est le codage en fréquence. Le principe est de caractériser chaque point par une fréquence de résonance différente de celle du point voisin. Pour que les noyaux précessent à des fréquences différentes en fonction de leur position, il suffit de créer volontairement des variations dans le champ magnétique  $\vec{B}_0$ . Ainsi un signal de fréquence différente sera attribué à chaque point de l'échantillon.

## B. Gradient de champ magnétique

Pour obtenir un champ magnétique  $\vec{B}_0$  non-uniforme, une technique simple est de le faire varier de façon linéaire dans une direction donnée. Prenons l'exemple d'un champ  $\vec{B}_0$  décroissant de haut en bas. La fréquence de résonance des points situés en haut de l'échantillon sera supérieure à celle des points plus bas, et tous les points se trouvant à une même hauteur résonneront à la même fréquence. Dans ce cas, le gradient permet donc de définir des plans perpendiculaires à la hauteur. Plus généralement, il est possible de définir des plans d'égale fréquence selon la direction du gradient.

### 3.3.4. Agents de contraste

Le contraste d'une image RMN dépend des différences de densité en hydrogène dans les tissus étudiés. Cependant, ces différences sont naturellement faibles. Pour améliorer le contraste de l'image, une technique consiste à activer spécifiquement les noyaux d'hydrogène des zones que l'on souhaite analyser, en injectant au sujet étudié un agent de contraste. En IRM, les agents de contrastes sont des molécules injectées *in vivo* (généralement par voie intraveineuse), qui diminuent le temps de relaxation des protons, c'est-à-dire le temps nécessaire aux noyaux pour retrouver leur état d'équilibre. Les temps de relaxation entre les tissus dans lesquels l'agent s'est distribué et ceux où il n'est pas distribué, seront donc différents, ce qui conduira à augmenter le contraste de l'image entre ces deux groupes de tissus. De plus, les agents de contraste permettent de réduire le temps d'acquisition d'une mesure, donc d'accumuler plus de mesures en un temps donné, augmentant ainsi la sensibilité et le contraste de l'image.

### 3.3.5. La place de l'IRM dans la recherche pharmaceutique

Les premiers développements en imagerie datent du début des années 1970. Les premières images chez l'homme ont été réalisées à la fin des années 1970. L'évolution de ces techniques a donc été très rapide. Aujourd'hui, l'IRM est devenue une technique majeure de l'imagerie médicale moderne, et est notamment développée dans le suivi des maladies chroniques, et dans la détection de pathologies au niveau tissulaire (Rudin *et al.*, 1999).

Grâce à son caractère non-invasif, l'IRM peut intervenir dans le suivi des maladies chroniques, qui ont tendance à augmenter avec l'âge de la population. Comme exemple, on peut citer les maladies neurodégénératives (telles que les maladies d'Alzheimer et Parkinson), les dégénérescences articulaires (telles que la décalcification), ou encore les inflammations chroniques (comme les tendinites). L'IRM offre la possibilité de suivre l'évolution de la maladie chez un même sujet, et aussi de évaluer les réponses à un traitement thérapeutique (Beckmann *et al.*, 2001). Un des domaines qui a le plus profité de l'essor des techniques d'imagerie est l'étude du cerveau (Aine, 1995; Berns, 1999). En effet, de nombreux groupes travaillent sur des méthodes permettant de relier les zones activées du cerveau lors de l'exécution d'une tâche ou lors d'une question précise (Ardekani and Kanno, 1998; Everitt and Bullmore, 1999; Genovese, 2000; Gossel *et al.*, 2001; Lazar *et al.*, 2001; Rabe-Hesketh *et al.*, 1997). Ces travaux ont directement des

applications fondamentales en sciences cognitives, telles que la neurophysiologie de la cognition. De plus, avec la détermination de ces zones, il serait possible d'apporter des diagnostics et de suivre des réponses thérapeutiques individualisées pour chaque patient.

Outre des informations purement anatomiques, l'IRM permet de fournir des renseignements sur les tissus, tels que leur perfusion. Plusieurs techniques d'analyse d'image ont été mises au point pour permettre la détection et la caractérisation (vasculaire, par exemple) de tumeurs de manière non-invasive (Brix *et al.*, 1991; Evelhoch, 1999; Evelhoch *et al.*, 1992; Okuhata *et al.*, 1999; Taylor and Reddick, 2000; Taylor *et al.*, 1999; Tofts, 1997; Tofts *et al.*, 1995; Tofts and Berkowitz, 1994; Tofts *et al.*, 1999; Tofts and Kermode, 1991) ou encore d'apporter un diagnostic sur le caractère malin ou bénin des tumeurs (Ando *et al.*, 2000; Brasch *et al.*, 1998; Hayton *et al.*, 1997). Là encore, l'IRM pourra être utilisée comme outil de suivi pour l'évolution des tumeurs, et pour évaluer les effets d'un traitement.

### 3.4. L'agrégation de système d'équations différentielles

Dans cette section, nous introduisons des méthodes d'agrégation ayant pour objectif de réduire un système d'équations différentielles, tel qu'un modèle PBPK, en un système de dimension moindre en regroupant (ou agrégeant) certaines variables d'état du système original entre elles. La manipulation de modèles à grande dimension (tels que certains modèles PBPK) peut en effet s'avérer difficile dans le cadre d'études d'optimisation d'un protocole expérimental, par exemple. Dans ces conditions, il peut être nécessaire de réduire les modèles. Nestorov *et al.* (1998) ont proposé une méthodologie permettant de regrouper certains compartiments d'un même modèle PBPK. Les regroupements de compartiments dépendent de la distribution de la substance en leur seins, et de la position des compartiments considérés dans la structure du modèle. Par exemple, des tissus agrégés doivent avoir des constantes de temps similaires ou très proches. L'agrégation du modèle dépendra donc essentiellement des valeurs des paramètres prises par la substance d'intérêt. En parallèle, des méthodes mathématiques d'agrégation ont été développées (Kuo and Wei, 1969; Li and Rabitz, 1989, 1990; Wei and Kuo, 1969), et appliquées avec succès dans le domaine de la combustion et de la chimie pétrolière (Li and Rabitz, 1993; Verhaar *et al.*, 1997; Wang *et al.*, 1998). Ces techniques d'agrégation sont basées uniquement sur la structure du modèle, et sont donc généralisables à de nombreux modèles. De plus, elles peuvent être appliquées lorsque plusieurs modèles sont considérés simultanément, par exemple lors de l'étude des interactions au sein de mélanges de substances.

Dans cette section, nous présentons l'agrégation mathématique de système d'équations différentielles, développée par Li et Rabitz (1989; 1990; 1991a; 1991c). Dans une section ultérieure, nous appliquerons ces techniques afin de réduire des modèles toxico/pharmacocinétiques. Dans cette présentation, nous nous limiterons donc aux systèmes d'équations différentielles linéaires ayant des coefficients d'échanges, entre les variables d'état, constants dans le temps.

#### 3.4.1. L'agrégation linéaire

Nous présentons les techniques développées par Li et Rabitz (1989; 1990) pour réduire les systèmes d'équations différentielles linéaires. Dans un premier temps, l'agrégation exacte d'un tel système est présentée. Ensuite, nous nous intéressons à l'agrégation approximative qui se révèle indispensable lorsque que l'on désire imposer des contraintes sur le nouveau système agrégé.

##### A. Agrégation linéaire exacte

Considérons un système d'équations différentielles décrivant les cinétiques de  $n$  variables d'état n'interagissant pas entre elles (par exemple, les concentrations ou quantités d'un modèle pharmacocinétique). Un tel système peut s'écrire

$$\frac{\partial \mathbf{y}(t)}{\partial t} = \mathbf{K}\mathbf{y}(t) + \mathbf{u}(t) \quad (2.13)$$



où  $\mathbf{y}$  est le vecteur (de dimension  $n$ ) contenant les  $n$  variables d'état du système,  $\mathbf{K}$  la matrice constante contenant les coefficients d'échanges entre les variables d'état (de dimension  $n \times n$ ), et  $\mathbf{u}$  un vecteur d'entrées (de dimension  $n$ ). On dit que le système (2.13) est exactement agrégeable par une matrice constante  $\mathbf{M}$  (de dimension  $\hat{n} \times n$ ) de rang  $\hat{n}$  ( $\hat{n} \leq n$ ), si pour

$$\hat{\mathbf{y}}(t) = \mathbf{M}\mathbf{y}(t) \quad (2.14)$$

il est possible de trouver une matrice  $\hat{\mathbf{K}}$  (de dimension  $\hat{n} \times \hat{n}$ ), et un vecteur d'entrée  $\hat{\mathbf{u}}$  tels que

$$\frac{\partial \hat{\mathbf{y}}(t)}{\partial t} = \hat{\mathbf{K}}\hat{\mathbf{y}}(t) + \hat{\mathbf{u}}(t) \quad (2.15)$$

Cependant tous les systèmes de la forme de (2.13) ne sont pas exactement agrégeables. Li et Rabitz (1989) ont montré qu'une condition nécessaire et suffisante à l'agrégation exacte est

$$\hat{\mathbf{K}}\hat{\mathbf{y}}(t) = \mathbf{M}\mathbf{K}\overline{\mathbf{M}}\mathbf{y}(t) \quad (2.16)$$

où  $\overline{\mathbf{M}}$  est un inverse généralisé de  $\mathbf{M}$ , défini par

$$\mathbf{M}\overline{\mathbf{M}} = \mathbf{I}_{\hat{n}} \quad (2.17)$$

avec  $\mathbf{I}_{\hat{n}}$  la matrice identité de rang  $\hat{n}$ . Il n'est cependant pas possible d'appliquer directement l'équation (2.16), puisque que la matrice  $\mathbf{M}$  est inconnue. Notons  $\mathbf{J}(\mathbf{y})$  la matrice jacobienne de  $\mathbf{K}\mathbf{y}(t)$ . Pour obtenir  $\mathbf{M}$ , deux étapes sont nécessaires (Li and Rabitz, 1989) : (i) déterminer si les sous-espaces  $\mathcal{M}$  invariants fixés non-triviaux de  $\mathbf{J}^T(\mathbf{y})$  existent; (ii) s'ils existent, examiner s'ils satisfont l'équation (2.16). Quand ces deux conditions sont remplies, le système décrit par l'équation (2.13) est exactement agrégeable par une matrice  $\mathbf{M}$ , dont les lignes sont composées des vecteurs de base de  $\mathcal{M}$ .

Pour construire la matrice  $\mathbf{M}$ , il reste donc à déterminer les sous-espaces  $\mathcal{M}$  invariants fixés non-triviaux de  $\mathbf{J}^T(\mathbf{y})$ . La transposée de la matrice jacobienne,  $\mathbf{J}^T(\mathbf{y})$ , du système original (2.13) peut être décomposée en une combinaison linéaire de matrices constantes  $\mathbf{A}_k$  appropriées, c'est-à-dire

$$\mathbf{J}^T(\mathbf{y}) = \sum_{k=1}^m a_k(\mathbf{y})\mathbf{A}_k \quad (2.18)$$

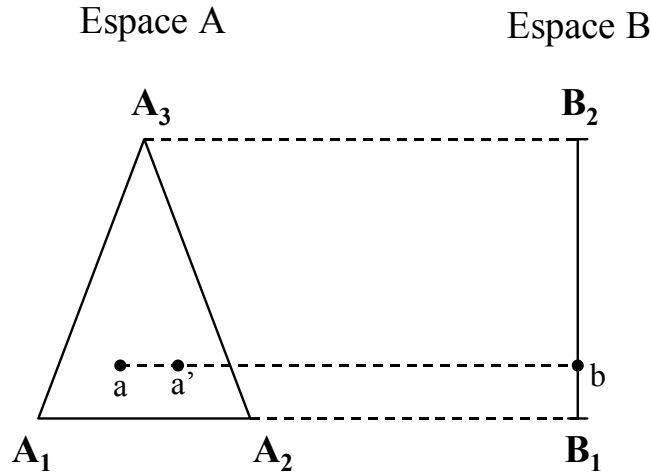
où  $m$  est plus petit que  $n^2$ , et les matrices constantes  $\mathbf{A}_k$  sont vues comme un ensemble de matrices de base de  $\mathbf{J}^T(\mathbf{y})$ .

Dans le cadre d'un système linéaire, la matrice  $\mathbf{J}^T(\mathbf{y})$  est constante. Dans cette situation particulière, les sous-espaces  $\mathcal{M}$  invariants fixés existent et correspondent à des valeurs propres constantes. La matrice jacobienne d'un système de réactions unimoléculaires, tel que (2.13), est juste  $\mathbf{K}$ . On a donc

$$\mathbf{J}^T(\mathbf{y}) = \mathbf{K}^T \quad (2.19)$$

Comme tous les sous-espaces invariants de  $\mathbf{J}^T(\mathbf{y})$  peuvent être utilisés pour construire une matrice d'agrégation, nous avons uniquement besoin de déterminer les sous-espaces  $\mathbf{K}^T$ -invariants. En considérant que les valeurs propres de  $\mathbf{K}^T$  sont distinctes, chaque sous-espace engendré par un sous-ensemble de ces vecteurs propres lui est invariant. La matrice d'agrégation  $\mathbf{M}$  sera donc composée des vecteurs propres de  $\mathbf{K}^T$ .

L'agrégation peut aussi être vue comme la projection du vecteur  $y$  (variables d'état du système original) dans un espace de dimension  $\hat{n}$ . La Figure 2.11 illustre la projection d'un espace A de dimension 3 dans un espace B de dimension 2. On remarque que deux vecteurs distincts dans A ( $a$  et  $a'$ ) peuvent correspondre au même vecteur dans B. Si  $\mathbf{M}$  est la matrice de projection de A dans B, on dit alors que  $a$  et  $a'$  sont  $\mathbf{M}$ -équivalents.



**Figure 2.11 :** Agrégation d'un espace A de dimension 3 en un espace B de dimension 2

### B. Agrégation linéaire contrainte : la méthode DCAL

En pratique, il est souvent utile de poser des conditions sur le système réduit afin que ce dernier puisse répondre aux questions d'intérêt (telles que déterminer la quantité d'une substance dans un compartiment précis). Dans ce but, il peut être nécessaire de laisser certaines variables d'état non-agrégées, par exemple. Pour tenir compte de ces restrictions, des méthodes d'agrégation contrainte ont été développées (Li and Rabitz, 1990, 1991a, c). L'agrégation est contrainte par une spécification *a priori* d'une partie de la matrice d'agrégation,  $\mathbf{M}$  qui, est représentée par

$$\mathbf{M} = \begin{pmatrix} \mathbf{M}_G \\ \mathbf{M}_D \end{pmatrix} \quad (2.20)$$

où  $\mathbf{M}_G$  est donnée, et  $\mathbf{M}_D$  est à déterminer. Sous ces contraintes, l'agrégation ne peut pas être exacte et des techniques d'agrégation approximative doivent être développées. Dans une agrégation approximative, les lignes de la matrice  $\mathbf{M}$  ne sont pas invariantes par rapport à  $\mathbf{J}^T(\mathbf{y})$ . Li et Rabitz (1990) ont développé une méthode itérative permettant de déterminer une matrice globalement optimale, telle que l'espace engendré par les lignes de  $\mathbf{M}$  soit "presque"  $\mathbf{J}^T(\mathbf{y})$ -invariant, et ont montré que le choix optimal de  $\mathbf{M}$  pour une matrice d'agrégation orthonormale est simplement  $\mathbf{M}^T$ . Plus récemment, Li et Rabitz (Li and Rabitz, 1991a, c, 1993) ont proposé une méthode plus directe : l'agrégation directe approximative contrainte (DCAL pour Direct Constrained Approximate Lumping). La méthode DCAL est basée sur la détermination numérique des

matrices de base  $\mathbf{A}_k$  de l'équation (2.18), pour  $\mathbf{J}^T(\mathbf{y})$ , en utilisant les valeurs de  $\mathbf{y}$  dans la région d'intérêt de l'espace de composition de dimension  $n$ . Il a été montré que  $\mathbf{M}$  peut être déterminé par

$$\mathbf{X}^T \mathbf{M}^T = \mathbf{0} \quad (2.21)$$

où  $\mathbf{X}$  est "presque" orthogonal à

$$\mathbf{Z} = \begin{pmatrix} \mathbf{M}_G \\ \mathbf{M}_G \mathbf{A}_1^T \\ \mathbf{M}_G \left( \mathbf{A}_1^T \right)^{s_1-1} \\ \dots \\ \mathbf{M}_G \\ \mathbf{M}_G \mathbf{A}_m^T \\ \mathbf{M}_G \left( \mathbf{A}_m^T \right)^{s_m-1} \end{pmatrix} \quad (2.22)$$

où  $s_m$  est le rang de  $\mathbf{A}_k$  (ou fixé à  $n$ ). La procédure suivante permet de déterminer  $\mathbf{M}$  :

- transformer  $\mathbf{M}_G(\mathbf{A}_k^T)^i$  ( $i$  de 0 à  $\text{Rang}(\mathbf{A}_k) - 1$ ) en une matrice orthogonale  $\mathbf{Q}(G)_{ki}^T$  par l'orthogonalisation de Gram-Schmidt (Bellman, 1960),
- construire la matrice symétrique  $\mathbf{Y}$  définie par :

$$\mathbf{Y} = \sum_{k=1}^m \sum_{i=0}^{s_k-1} \mathbf{Q}(G)_{ki}^T \mathbf{Q}(G)_{ki} \quad (2.23)$$

- déterminer les valeurs et vecteurs propres de  $\mathbf{Y}$ .

Pour un système agrégé de dimension  $\hat{n}$ , les vecteurs propres correspondant aux  $n - \hat{n}$  plus petites valeurs propres définissent la matrice  $\mathbf{X}$ , et les vecteurs propres correspondant aux  $\hat{n}$  valeurs propres de  $\mathbf{Y}$  définissent les lignes de  $\mathbf{M}$ . Pour forcer  $\mathbf{M}_G$  à correspondre aux vecteurs propres de  $\mathbf{Y}$  ayant les valeurs propres les plus grandes,  $\mathbf{M}_G$  dans  $\mathbf{Z}$ , équation (2.22), est multiplié par une grande constante. Ceci permet de s'assurer que les variables non agrégées spécifiées dans  $\mathbf{M}_G$  font partie du système agrégé. Lorsque les matrices de base  $\mathbf{A}_k$  de  $\mathbf{J}^T(\mathbf{y})$  sont utilisées, la matrice  $\mathbf{M}$  résultante est globalement optimale (Li and Rabitz, 1991a).

### 3.4.2. Exemple d'agrégation exacte d'un système de dimension 3

Dans cette partie, nous présentons un exemple d'application d'agrégation exacte. Nous considérons un modèle de dimension 3, et proposons de le réduire en un système de dimension 2. La dimension initiale du système ne permet pas une réduction plus importante. Cependant, cet exemple simple permet d'entrevoir les avantages de l'agrégation sur des systèmes à grandes dimensions (en industrie pétrolière, les modèles peuvent concerner plus de  $10^2 - 10^3$  substances). Des exemples d'application de l'agrégation contrainte linéaire seront considérés, dans la suite, sur des modèles toxico/pharmacocinétiques.

### A. Définition du système

Considérons le système de réactions chimiques suivant :



où  $Y_i$  représente la substance  $i$ , et  $k$  la constante de réaction (supposée unique pour toutes les réactions). Notons  $y_i$  la concentration de  $Y_i$ . Les réactions chimiques de (2.24) peuvent se traduire par le système suivant :

$$\begin{cases} \frac{\partial y_1(t)}{\partial t} = -ky_1(t) + ky_2(t) \\ \frac{\partial y_2(t)}{\partial t} = ky_1(t) - 2ky_2(t) + ky_3(t) \\ \frac{\partial y_3(t)}{\partial t} = ky_2(t) - ky_3(t) \end{cases} \quad (2.25)$$

Ce système est de la forme du système (2.13) avec le vecteur d'entrées  $\mathbf{u}$  égal à zéro, et la matrice des coefficients  $\mathbf{K}$ ,

$$\mathbf{K} = \begin{pmatrix} -k & k & 0 \\ k & -2k & k \\ 0 & k & -k \end{pmatrix} \quad (2.26)$$

### B. Les schémas d'agrégation

D'après la méthodologie d'agrégation linéaire exposée ci-dessus, la matrice d'agrégation  $\mathbf{M}$  est construite à partir des vecteurs propres de  $\mathbf{K}^T$ . Les valeurs propres sont égales à  $\{0; -3k; k\}$ , et les vecteurs propres correspondants à

$$x_1 = \{1; 1; 1\}, \quad x_2 = \{1; -2; 1\}, \quad x_3 = \{-1; 0; 1\} \quad (2.27)$$

Les schémas d'agrégation se définissent à l'aide de l'espace et des sous-espaces engendrés par les vecteurs propres de  $\mathbf{K}^T$ . Le nombre de schémas possibles est alors égal à la dimension de l'espace engendré, c'est-à-dire  $2^n$  pour un système de dimension  $n$  ayant  $n$  valeurs propres distinctes. Dans notre cas, cet espace est défini par

$$\text{span}\{x_1, x_2, x_3\} = \{\{0\}, \{x_1\}, \{x_2\}, \{x_3\}, \{x_1, x_2\}, \{x_1, x_3\}, \{x_2, x_3\}, \{x_1, x_2, x_3\}\} \quad (2.28)$$

Supposons que nous souhaitons réduire le système original de dimension 3 en un système de dimension 2 ( $\hat{n} = 2$ ). Il suffit de choisir l'espace engendré par deux vecteurs propres, disons  $x_1$  et  $x_3$ . La matrice d'agrégation  $\mathbf{M}$  est donc

$$\mathbf{M} = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 0 & 1 \end{pmatrix} \quad (2.29)$$

L'inverse généralisé de  $\mathbf{M}$  est

$$\overline{\mathbf{M}} = \begin{pmatrix} \frac{1}{3} & -\frac{1}{2} \\ \frac{1}{3} & 0 \\ \frac{1}{3} & \frac{1}{2} \end{pmatrix} \quad (2.30)$$

Il est possible de multiplier  $\mathbf{M}$  par une matrice non-singulière  $\mathbf{N}$  (de dimension  $\hat{n} \times \hat{n}$ ), sans changer les propriétés d'agrégation de  $\mathbf{M}$ . Cette manipulation peut s'avérer utile, lorsque l'on souhaite se ramener à un système avec des coefficients positifs (quand les variables sont des concentrations, par exemple). Définissons une telle matrice  $\mathbf{N}$ ,

$$\mathbf{N} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad (2.31)$$

La nouvelle matrice d'agrégation est obtenue par

$$\mathbf{M}_1 = \mathbf{N} \times \mathbf{M} = \begin{pmatrix} 0 & 1 & 2 \\ 2 & 1 & 0 \end{pmatrix} \quad (2.32)$$

dont l'inverse généralisé est

$$\overline{\mathbf{M}}_1 = \overline{\mathbf{M}} \times \overline{\mathbf{N}} = \begin{pmatrix} -\frac{1}{12} & \frac{5}{12} \\ \frac{1}{6} & \frac{1}{6} \\ \frac{5}{12} & -\frac{1}{12} \end{pmatrix} \quad (2.33)$$

### C. Formulation du système agrégé

Notons  $\hat{\mathbf{y}}$  le vecteur contenant les variables du système réduit, et  $\hat{\mathbf{K}}$  la matrice des coefficients correspondante. D'après l'équation (2.16), on a la relation suivante

$$\hat{\mathbf{K}}\hat{\mathbf{y}}(t) = \mathbf{M}_1 \mathbf{K} \overline{\mathbf{M}}_1 \hat{\mathbf{y}}(t) \quad (2.34)$$

Il est ensuite facile d'obtenir la matrice  $\hat{\mathbf{K}}$

$$\hat{\mathbf{K}} = \frac{1}{2} \times \begin{pmatrix} -k & k \\ k & -k \end{pmatrix} \quad (2.35)$$

Le système réduit est donc

$$\hat{\mathbf{Y}}_1 \xleftrightarrow{\frac{k}{2}} \hat{\mathbf{Y}}_2 \quad (2.36)$$

dont les nouvelles variables d'état  $(\hat{y}_1, \hat{y}_2)$  sont des combinaisons linéaires des variables originales

$$\hat{\mathbf{y}} = \begin{pmatrix} y_2 + 2y_3 \\ 2y_1 + y_2 \end{pmatrix} \quad (2.37)$$

### 3.5. Objectifs spécifiques

Typiquement, les seules données toxicocinétiques disponibles sont la concentration de la substance dans certains fluides biologiques, tels que le sang ou l'urine. Dans des études animales, il est possible d'obtenir la concentration interne tissulaire à un temps donné, avec sacrifice de l'animal exposé (ce qui limite en fait le nombre d'études de ce type). Pour les études conduites chez l'humain, la concentration interne des tissus (ou compartiments) est non-observée, et doit être inférée. Les données disponibles pour calibrer le modèle toxicocinétique sont donc généralement peu nombreuses et éparpillées. Cette pauvreté des données peut entraîner la non-identifiabilité de certains paramètres du modèle dans certains cas (par exemple, pour un modèle de grande dimension). Pour remédier à ce problème, nous proposons de développer deux protocoles expérimentaux. Dans ce travail, nous souhaitons développer de nouveaux outils et concepts permettant de paramétrer les modèles PBPK. Deux axes de recherche sont proposés : la collecte de données sur la distribution spatio-temporelle d'une substance dans l'organisme par des techniques d'imagerie, et l'utilisation de traceurs non-toxiques lors d'études toxicocinétiques.

Une première solution est de développer des techniques permettant de quantifier la présence d'une substance au niveau tissulaire, et ensuite d'utiliser ces données pour calibrer le modèle. Cette quantification est actuellement réalisable pour certaines substances à l'aide de techniques d'imagerie quantitatives non invasives, telles que l'imagerie par résonance magnétique (IRM) ou la tomographie par émission de positons (TEP). De plus, les images obtenues à l'aide de produits de contraste contiennent de nombreuses informations sur l'anatomie et la physiologie du sujet étudié, qui pourraient être utilisées pour l'estimation des paramètres de modèles physiologiques. Dans ce travail, nous développons donc la modélisation PBPK pour décrire les données de cinétiques obtenues par des techniques d'imagerie, et en particulier par IRM.

Notre second axe de recherche est de développer l'utilisation des traceurs non-toxiques dans les études toxicocinétiques. Nous proposons des protocoles complémentaires d'exposition à une substance chimique et à un traceur non-toxique. L'idée de base d'un tel couplage est, dans un premier temps, de paramétrer partiellement le modèle PBPK à l'aide de nombreuses données expérimentales sur le traceur, puis de compléter cette paramétrisation à l'aide d'un minimum de données sur la substance toxique d'intérêt. Notre hypothèse est que les données sur le traceur permettent une détermination efficace des caractéristiques physiologiques propres de chaque sujet étudié. Pour chaque exposition, il sera donc nécessaire d'analyser les deux ensembles de données (toxique et traceur). Les calculs induits par ce type d'analyse étant lourds, il est particulièrement intéressant d'analyser les données après réduction du système d'équations différentielles décrivant le modèle (Toth *et al.*, 1997). L'essentiel de ces méthodes de réduction est de préserver la dynamique des systèmes tout en diminuant le nombre des variables d'état. Ces méthodes seront appliquées sur quelques modèles PBPK afin d'évaluer leur potentiel dans ce domaine.

Les analyses effectuées pour ces deux protocoles sont conduites dans un cadre statistique bayésien. Chaque modèle PBPK développé est calibré à l'aide de données expérimentales par des méthodes de

simulation Monte Carlo par chaîne de Markov. Nous avons choisi ce cadre inférentiel pour deux raisons principales. La première concerne la possibilité de définir des distributions *a priori* pour tous les paramètres du modèle, et notamment des distributions informatives pour les paramètres physiologiques dont les valeurs sont relativement bien connues. La seconde raison est l'estimation de l'incertitude contenue dans les paramètres, grâce à l'obtention d'une distribution de probabilité *a posteriori* pour chacun des paramètres.

## **4. L'analyse statistique bayésienne de données toxicocinétiques**



# **L'Analyse Statistique Bayésienne de Données Toxicocinétiques**

## **Bayesian Statistical Analysis of Toxicokinetic Data**

Sandrine Micallef<sup>1</sup>, Céline Brochot<sup>1</sup> et Frédéric Y. Bois<sup>1</sup>

<sup>1</sup>INERIS, Institut National de l'Environnement Industriel et des Risques, Unité de Toxicologie Expérimentale, Parc Alata BP2, 60550 Verneuil En Halatte, France

Soumis à Environnement, Risque et Santé

## RESUME

L'évaluation de la toxicité des produits chimiques passe par une étude de leur toxicocinétique. Cette dernière vise à établir un lien entre l'exposition à une substance toxique et la quantité de celle-ci atteignant les tissus cibles de l'action toxique. Pour ce faire, l'expérimentation recueille des données de distribution spatiale et temporelle de la substance dans l'organisme étudié (par exemple, des mesures de concentration sanguine en différents instants). Il est ensuite possible de construire des modèles mathématiques paramétrés (modèles toxicocinétiques) pour caractériser la relation entre l'exposition et la quantité interne de toxique. Ces modèles sont généralement des modèles compartimentaux et reposent fréquemment sur une description physiologique de l'organisme. Des traitements statistiques appropriés permettent d'ajuster ces modèles à l'aide des données expérimentales. Nous décrivons comment l'analyse bayésienne permet de prendre en compte et d'estimer l'incertitude et la variabilité inhérentes aux données toxicocinétiques. Elle intègre l'information disponible *a priori* sur la valeur des paramètres, limitant, de ce fait, le nombre d'expositions expérimentales nécessaires. Le processus bayésien d'analyse de données comprend l'estimation des paramètres (calibration), la vérification et la validation des modèles. Ce processus peut ensuite être complété par le choix entre plusieurs modèles construits à partir d'hypothèses différentes, l'optimisation de nouveaux protocoles expérimentaux. Une illustration de l'approche bayésienne est proposée pour le cas du 1,3-butadiène, substance chimique considérée comme potentiellement carcinogène pour l'humain.

Mots-Clés : Modèles toxicocinétiques; Statistique bayésienne; Simulations stochastiques; 1,3-Butadiène

## SUMMARY

Risk assessment for chemicals calls for toxicokinetic study to establish a link between exposure and the quantity reaching target tissues or cells. To characterize this link, experimental data on the spatiotemporal distribution of the substance in the body (*e.g.*, blood concentrations at various times) can be collected, and analyzed with parametric models (called toxicokinetic, TK, models). These models are generally compartmental, and can be based on physiology (PBTK models). Adequate statistical treatment allows fitting of TK/PBTK models to the data. Bayesian analysis, in particular, takes into account and estimates uncertainty and variability inherent to TK data. In that framework, *prior* information on parameter values can be integrated in the estimation procedure, limiting the need for experimental exposures. In this paper, we detail the Bayesian process of TK analysis, including parameter estimation (or calibration), checking and validation of models. The process can be completed with model choice between competing models, or with the optimal design of experimental protocols. For an illustration, we analyze and model the toxicokinetics of 1,3-butadiene, a potential carcinogen for humans.

Keywords: Toxicokinetic modeling; Bayesian statistics; Stochastic simulations; 1,3-Butadiene

## INTRODUCTION

Les études toxicologiques ont pour objectif de comprendre et quantifier le lien entre l'exposition à une substance chimique et les effets qui en découlent. L'établissement de ce lien rend possible l'évaluation et la prédiction des risques encourus par les populations exposées à un agent toxique. La représentation du mécanisme d'intoxication ressort de deux disciplines : la toxicocinétique (description spatiale et temporelle de la concentration de la substance dans l'organisme) et la toxicodynamie (détermination de la relation entre la concentration interne et l'effet résultant). Cet article fait le point sur l'analyse statistique de données *toxicocinétiques*. Il prolonge la revue de la question faite par Bernillon et Bois (2000), traitant essentiellement de la spécification et de la calibration bayésienne des modèles.

Les données disponibles au moment de répondre à une question toxicologique peuvent être classées en trois catégories : les résultats d'expérimentations anciennes ou réalisées spécifiquement pour répondre à la question, les covariables et l'information *a priori*.

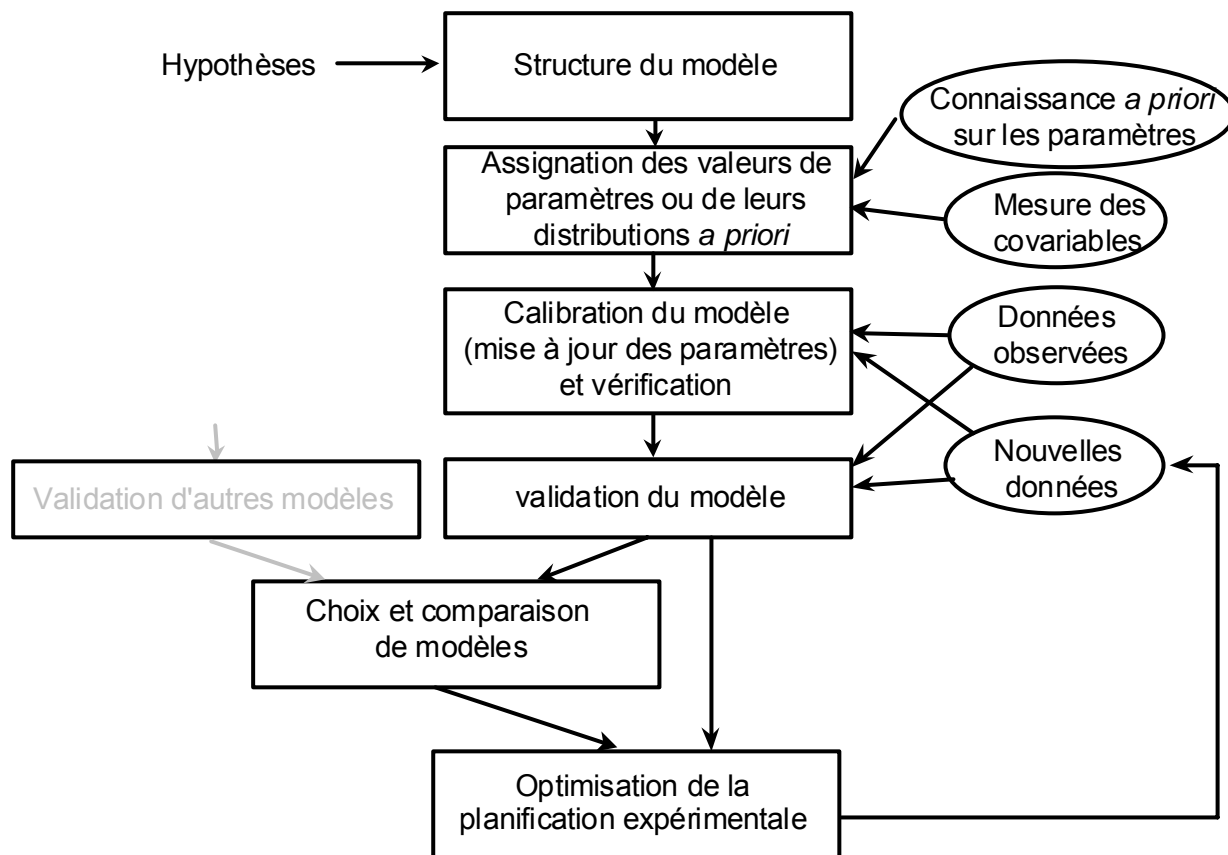
- Les données expérimentales toxicocinétiques consistent en des mesures de concentrations ou de quantités de substance relevées en divers tissus ou fluides biologiques, au cours du temps, chez des animaux ou des volontaires humains exposés (voir Figure 2). Elles peuvent avoir été collectées lors d'études précédentes ou pour répondre à la question posée. L'expérimentation, si elle est nécessaire, doit respecter des contraintes éthiques et financières, qui font que les données toxicologiques sont assez éparpillées et obtenues sur une partie de la population concernée.
- Les covariables sont des caractéristiques directement mesurables sur les sujets avec une erreur négligeable. Typiquement, elles incluent l'âge, la masse corporelle, et le sexe des sujets considérés.
- Pour d'autres caractéristiques physiologiques (tels que le volume des organes ou les flux sanguins), des intervalles de valeurs plausibles sont souvent disponibles dans la littérature (Brown *et al.*, 1997; International Commission on Radiological Protection (ICRP), 2002). Ces informations, dites "*a priori*", proviennent d'observations ou d'expérimentations *in vitro* ou *in vivo* antérieures et constituent une large part du corpus des connaissances en toxicologie.

Le rôle des covariables et des informations *a priori* diffère de celui des données expérimentales toxicocinétiques. Alors que les premières renseignent plus ou moins directement sur la valeur des caractéristiques spécifiques des sujets, les données expérimentales informent indirectement sur les paramètres de la biodistribution de l'agent considéré. La volonté, ou même la nécessité scientifique, de prendre en compte l'ensemble des informations disponibles et leurs incertitudes amènent à analyser les données toxicocinétiques dans un contexte statistique bayésien (Bernardo and Smith, 1994; Gelman *et al.*, 1995; Wakefield and Bennett, 1996). Un tel traitement ne peut se passer de modèles mathématiques.

Les modèles toxicocinétiques sont des outils mathématiques capables de décrire la distribution locale et temporelle d'un agent dans un organisme en fonction de ses caractéristiques (Bois, 2003; Rowland *et al.*, 1973; Tozer, 1981). Ils peuvent donc faire le lien entre covariables, données *a priori* et données expérimentales. Ces modèles simplifient la complexité physiologique en divisant le corps en

"compartiments" et en considérant que les substances se distribuent uniformément à l'intérieur de chaque compartiment (Gibaldi and Perrier, 1982). Pour autant que la structure anatomo-physiologique du corps soit sensiblement la même pour les individus d'une même espèce, de tels modèles sont capables de décrire la toxicocinétique d'une substance pour tous les sujets d'une population. Les différences entre individus peuvent ensuite s'expliquer par des valeurs de paramètres différentes d'un sujet à l'autre.

Un modèle adapté à l'analyse de données toxicocinétique se construit en plusieurs étapes. Dans un premier temps, la formulation d'hypothèses amène à structurer le modèle global et à en définir le domaine de validité. L'étape suivante consiste à construire un modèle détaillant le processus de distribution dans le corps. La modélisation doit ensuite être complétée par un modèle statistique tenant compte des incertitudes de mesure et/ou de la variabilité entre les observations de différents individus. Dans un cadre bayésien, une fois le modèle global construit, les informations disponibles *a priori* sont rassemblées pour assigner des distributions de probabilité aux paramètres. Ces distributions sont ensuite mises à jour à partir des données observées. Cette étape constitue la "calibration" du modèle. Les étapes suivantes concernent la vérification de l'adéquation du modèle (en comparant ses prédictions sur le jeu de données ayant servi à calibrer le modèle, par exemple), sa validation, et éventuellement, la redéfinition de son domaine de validité. Si plusieurs modèles ont été définis et validés, il est possible de les comparer en mettant en place une procédure de choix de modèle. Enfin, le modèle choisi peut être inclus dans une procédure d'optimisation d'un plan expérimental pour permettre de collecter de nouvelles données de façon optimale. Par la suite, les nouvelles données peuvent être reprises dans le processus et permettre une nouvelle calibration. Ce processus global est résumé Figure 1. Cet article, explicite, dans le détail, ces différentes étapes.



**Figure 1 :** Schéma du processus de construction d'un modèle dans le cadre bayésien. Les cadres rectangulaires représentent chacun des pas du processus et les formes ovales indiquent les différents types de données utilisées. Le rectangle grisé correspond à une étape d'un processus parallèle.

## QUESTION TOXICOLOGIQUE ET DONNEES EXPERIMENTALES

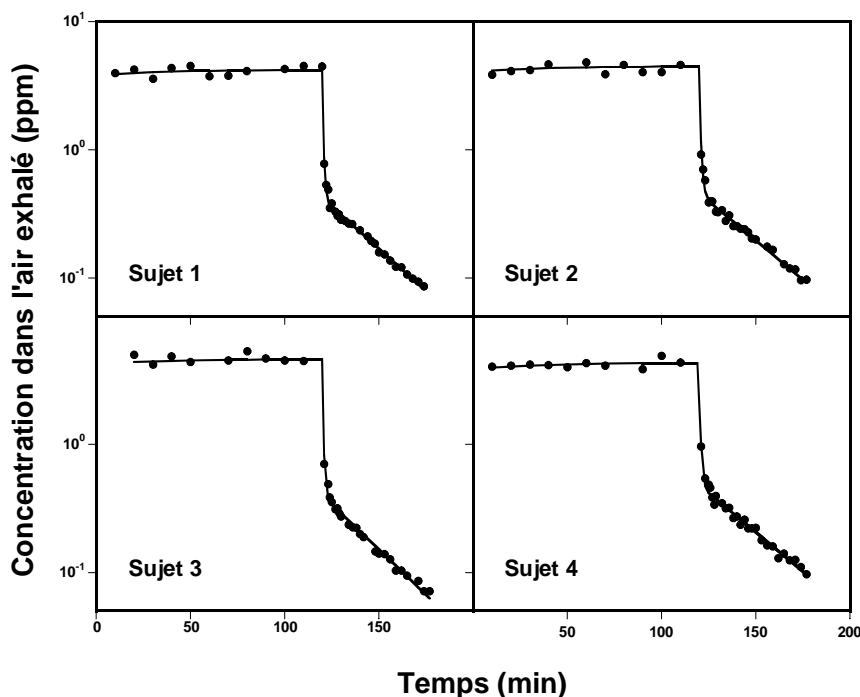
Pour illustrer le processus bayésien de construction d'un modèle toxicocinétique, nous prendrons le cas de l'estimation du métabolisme du 1,3-butadiène dans une population humaine. Le 1,3-butadiène est une substance chimique produite, notamment, lors de la combustion de produits pétroliers ou encore dans la manufacture des caoutchoucs, des latex et des résines synthétiques (International Agency for Research on Cancer (IARC), 1992). Il pénètre généralement dans l'environnement par les gaz d'échappement des véhicules et les gaz d'émission des foyers de combustion. Environ 70 000 travailleurs des industries pétrochimiques et plastiques y sont potentiellement exposés. Une exposition aiguë (>1000 parties par million en volume, ppm) au butadiène peut entraîner des dommages au niveau du système nerveux central, des troubles visuels, des nausées ou encore une baisse de la pression sanguine. En outre, des effets cancérogènes seraient potentiellement causés par l'exposition chronique à cette substance. Les effets cancérogènes ne

seraient pas engendrés par le butadiène lui-même mais par ses métabolites (produits de sa dégradation enzymatique). Pour évaluer les risques encourus après une exposition au butadiène, sous l'hypothèse qu'il est indirectement cancérigène, il est donc nécessaire de quantifier son métabolisme. Chez l'homme, une telle quantification est difficile puisqu'elle doit être réalisée à partir d'observations indirectes (par exemple, à partir de concentrations de butadiène mesurées dans l'air exhalé suite à de très faibles expositions). Nous proposons de l'estimer par l'intermédiaire d'une modélisation de sa toxicocinétique dans l'organisme humain.

Dans une première étude (étude I), huit volontaires humains ont été recrutés pour être exposés à de faibles concentrations de butadiène (correspondant à une journée d'exposition professionnelle autorisée par l'Occupational Safety and Health Administration des Etats Unis). Le protocole d'étude a été présenté et accepté par un comité d'éthique et chaque volontaire a reçu et signé un document de consentement éclairé. Ce protocole décrivait l'étude, ses objectifs et spécifiait les conditions d'expositions (durant 2 heures à une concentration de 5 ppm). Chaque sujet a passé une consultation médicale afin de vérifier l'absence de conditions pouvant altérer son métabolisme, mesurer sa masse corporelle, son débit de ventilation pulmonaire au repos et son coefficient de partage entre sang et air pour le butadiène (*in vitro* sur un prélèvement de sang). Les expositions par inhalation au butadiène ont ensuite été réalisées. Pour chaque sujet, environ 41 échantillons d'air exhalé ont été prélevés, 11 durant l'exposition et 30 dans les 57 minutes suivant celle-ci. Quelques sujets eurent un nombre de mesures inférieur à 41 (certains temps de mesure ayant été manqués), le plus petit nombre d'échantillons collectés pour un même sujet était 32. Les données obtenues ont été publiées dans Bois *et al.* (1999a). La Figure 2 présente les concentrations de butadiène dans l'air exhalé pour quatre sujets.

L'étude II avait pour objectif de déterminer si le métabolisme du butadiène dépendait de certaines covariables telles que le sexe, l'ethnie ou l'âge. Ces données ont été récoltées auprès de 138 sujets, en tirant partie (par optimisation formelle du plan expérimental) de l'information déjà obtenue grâce à la première exposition : seules 10 mesures ont été faites par sujet. Chaque exposition au butadiène était de 2 ppm pendant 20 minutes et les concentrations de butadiène dans l'air exhalé ont été relevées jusqu'à une heure après la fin de l'exposition.

Une dernière série de mesure (étude III) a été faite pour déterminer si le métabolisme des sujets variait au cours du temps. Onze sujets ont été exposés au butadiène à deux occasions pour étudier ces variations. A chaque occasion, le protocole d'exposition et de mesure de l'étude II a été appliqué.



**Figure 2 :** Concentrations de 1,3-butadiène exhalées suite à une exposition de 2 heures à 5 ppm. Les données présentées ont été relevées chez quatre volontaires différents et jusqu'à une heure après exposition. Les observations sont représentées par les points ; la ligne continue donne les prédictions du modèle à deux compartiments après ajustement.

## LE PROCESSUS D'ANALYSE DES DONNEES

### Modélisation

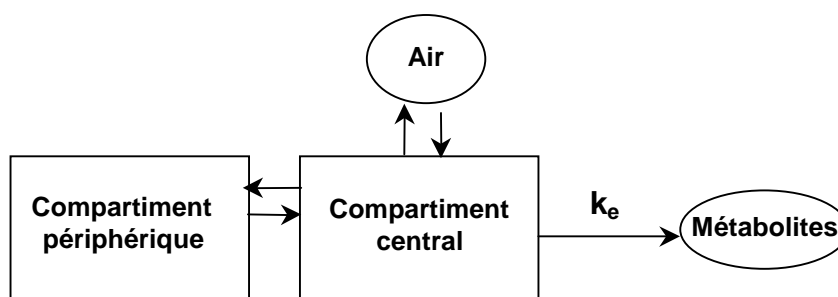
Dans le cas du butadiène, le modèle global inclut trois sous-modèles visant à caractériser les différents niveaux de variabilité et d'incertitude (Bernillon and Bois, 2000) : le modèle toxicocinétique explicitant le phénomène de biodistribution, le modèle de population représentant la variabilité au sein d'un groupe d'individus et le modèle décrivant l'erreur de mesure expérimentale.

#### Le modèle toxicocinétique

Pour décrire l'évolution de la concentration de butadiène dans l'air exhalé, un modèle toxicocinétique compartimental classique a été choisi (Figure 3). Deux compartiments ont été définis : un compartiment "central" par lequel ont lieu les échanges de butadiène avec le milieu extérieur (inhalation et élimination métabolique) et un compartiment "périphérique" connecté au compartiment central. Un système d'équations différentielles, donné en annexe, permet de décrire les variations temporelles des quantités de butadiène au sein de ces compartiments et sa cinétique d'élimination. Ce modèle a déjà été présenté dans Bois *et al.* (1999a). Il fait intervenir différents paramètres (physiologiques ou non), dont la constante de métabolisme ( $k_e$ ) qui nous intéresse plus particulièrement.



Chaque modèle n'est valide que sous certaines conditions. Il est donc nécessaire de définir le domaine de validité du modèle étant données les hypothèses admises lors de sa construction. Le modèle toxicocinétique à deux compartiments ne permet pas de décrire la phase terminale d'élimination contrôlée par la libération lente du butadiène contenu dans les graisses. Ce modèle n'est donc pas fiable pour décrire et prédire la distribution du butadiène lorsque la quantité libérée par les graisses ne peut plus être considérée comme négligeable. Étant donné que l'objectif de cette étude n'était pas de modéliser la phase terminale d'élimination, nous n'avons pas jugé nécessaire de complexifier le modèle par l'ajout d'un compartiment "graisses". Par ailleurs, ce modèle a été construit pour décrire des expositions à faibles doses. Un modèle simple du processus de métabolisme a donc pu être utilisé. L'élimination a été modélisée par une relation linéaire, uniquement valable pour les faibles doses. Ces deux hypothèses permettent ainsi de définir le domaine de validité du modèle toxicocinétique présenté.

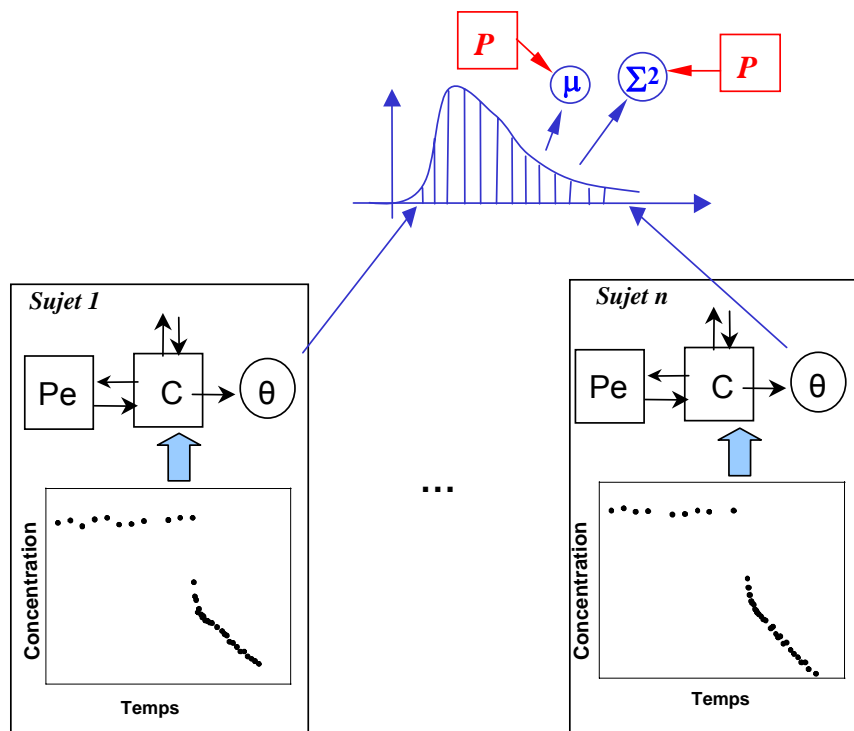


**Figure 3 :** *Modèle compartimental utilisé pour l'analyse des données de toxicocinétique du 1,3-butadiène. Ce modèle est constitué de deux compartiments ; les flèches caractérisent les échanges de matière.  $k_e$  est la constante de métabolisme.*

#### *Le modèle de population*

Une même substance se distribue différemment dans le corps de différents individus. Ceci peut s'expliquer par des variations des paramètres toxicocinétiques (tels que les constantes de transfert entre compartiments, la ventilation pulmonaire, la constante de métabolisme, les coefficients de partage *etc.*). Ces différences peuvent être prises en compte dans une analyse toxicocinétique à l'aide d'un modèle de population. Ces modèles ont été introduits initialement pour le développement et l'évaluation de substances médicamenteuses (Aarons, 1992; Beal and Sheiner, 1982; Sheiner and Beal, 1980; Wakefield and Bennett, 1996). Depuis, ils sont plus largement utilisés, notamment dans les études toxicocinétiques (Bois *et al.*, 1996b; Gelman *et al.*, 1996a). Leur objectif est d'estimer la variabilité au sein d'une large population à partir de données sur quelques individus. Ils supposent que les processus chimiques ou biologiques sous-jacents sont identiques pour chaque individu (même structure du modèle toxicocinétique) et que les différences observées s'expliquent par des variations aléatoires d'un sujet à l'autre pour les valeurs des paramètres toxicocinétiques. L'aléa pour ces paramètres peut être décrit par une loi de probabilité multivariée appelée "distribution de population". Les paramètres de cette loi sont dénommés paramètres de population.

Afin de tenir compte de la variabilité inter-individuelle, le modèle toxicocinétique à 2 compartiments présenté ci-dessus a donc été inséré dans un modèle hiérarchique de population (Figure 4). Dans la population, les paramètres toxicocinétiques individuels (de valeur *a priori* inconnue) sont supposés être distribués normalement (ou log-normalement) autour d'une moyenne  $\mu$  et d'une variance  $\Sigma^2$ , eux aussi de valeur inconnue.



**Figure 4 :** Schéma du modèle de population. Chaque bloc représente un sujet pour lequel les concentrations mesurées ont été représentées par des points. Pour chaque sujet, certains paramètres sont connus (covariables), par exemple, la masse corporelle ou le taux de ventilation (non représentés sur le graphe). Les paramètres inconnus, noté  $\theta$  (e.g., la constante de métabolisme  $k_e$ ) sont encadrés. Ces paramètres sont supposés distribués dans la population suivant une loi de moyenne  $\mu$  et de variance  $\Sigma^2$ . Dans le cadre bayésien, une distribution *a priori* ( $P$ ) est affectée à chacun des paramètres de population pour représenter la connaissance imparfaite que l'on en a avant l'expérimentation.

#### Le modèle d'erreur de mesure

Les données utilisées pour la calibration du modèle sont la plupart du temps des mesures (pour le butadiène, il s'agit de concentrations d'air exhalé). Elles sont donc naturellement entachées d'une erreur qu'il faut prendre en compte par un modèle d'erreur de mesure.

Pour des données biologiques strictement positives, le modèle d'erreur de mesure usuel est le modèle lognormal. Les paramètres de ce modèle (tels que la variance d'erreur de mesure) sont à estimer. Si des informations sont disponibles *a priori*, sur ces paramètres, par exemple à l'aide de données d'assurance qualité du laboratoire d'analyse, il est souhaitable d'en tenir compte. Les autres mesures individuelles

(ventilation pulmonaire, et coefficient de partage par exemple) peuvent être aussi entachées d'une incertitude de mesure.

### ***La calibration bayésienne***

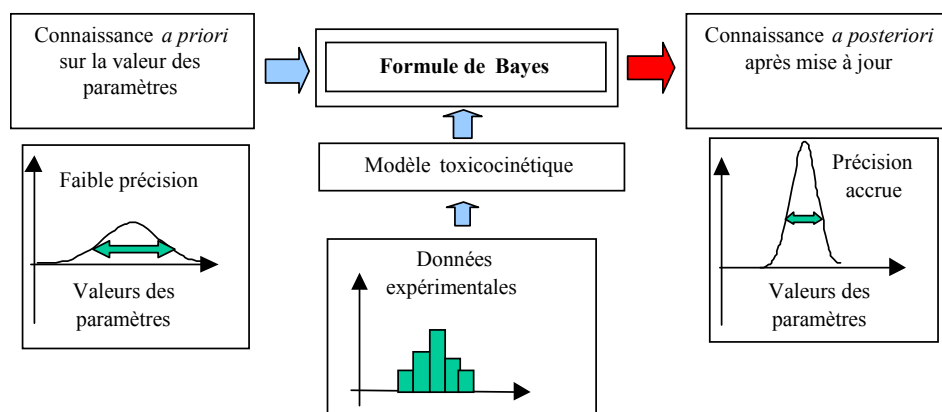
L'analyse bayésienne permet d'établir des distributions de probabilités pour les valeurs de paramètres ou de prédictions d'un modèle, à partir d'informations *a priori* et de données expérimentales. Une distribution de probabilité peut être définie comme une fonction associant les valeurs prises par une variable (dite aléatoire) à leur probabilité de réalisation. Dans cette section, nous commençons par présenter l'estimation bayésienne des paramètres d'un modèle à partir de données toxicocinétiques. Chaque étape du processus d'analyse de données est ensuite détaillée individuellement. L'exemple du butadiène est traité simultanément.

#### ***Le paradigme bayésien***

Dans un contexte bayésien, toutes les quantités observables ou non-observables (rassemblées sous la notation  $\theta$ ) sont considérées comme des variables aléatoires. La principale caractéristique de cette méthode est l'utilisation de distributions de probabilité pour quantifier les incertitudes sur les valeurs de ces variables. Ces distributions peuvent être interprétées en terme de degré de croyance sur les valeurs probables des paramètres et permettent, par exemple, de définir des intervalles de valeurs possibles. Pour chaque variable, une distribution de probabilité *a priori*,  $P(\theta)$ , est définie. Ces distributions sont ensuite mises à jour en utilisant de nouvelles données expérimentales (notées  $y$ ), conduisant à une distribution *a posteriori*,  $P(\theta|y)$ , (Figure 5). Cette mise à jour est effectuée par l'intermédiaire du théorème de Bayes, qui stipule que la distribution *a posteriori* des paramètres est proportionnelle au produit de toutes les distributions *a priori*  $P(\theta)$ , (sous l'hypothèse habituelle d'indépendance des paramètres), et de la vraisemblance des données  $P(y|\theta)$  :

$$P(\theta|y) = k \cdot P(y|\theta) \prod_{\theta} P(\theta)$$

où  $k$  est une constante. Des distributions *a priori* conjointes (non-indépendantes) peuvent également être utilisées. La forme de la vraisemblance dépend du modèle d'erreur et du modèle déterministe adopté. La distribution *a posteriori* représente la connaissance que l'on peut avoir sur les paramètres, y compris l'information *a priori* et celle contenue dans les données expérimentales. La détermination de cette loi *a posteriori* constitue l'estimation bayésienne des paramètres d'un modèle, encore appelée calibration. Il est souvent difficile d'obtenir une expression analytique pour la distribution *a posteriori* conjointe. Le développement des méthodes de Monte Carlo par chaînes de Markov, présentées dans la suite, a permis de contourner ces difficultés.



**Figure 5 :** Principe de la calibration bayésienne.

### *L'information a priori*

La calibration bayésienne d'un modèle passe par la définition d'une distribution *a priori* pour chacun de ses paramètres (observables ou non). Cette distribution doit refléter la connaissance sur ces paramètres, avant observation des données. Cette connaissance inclut, par exemple, des données *in vitro*, des informations référencées dans la littérature ou encore l'avis d'experts. La dispersion de la densité sera fonction de la connaissance sur le paramètre. Des distributions précises (informatives) sont définies pour les paramètres dont la valeur est assez bien connue, tels que certains paramètres physiologiques (flux sanguins ...). En revanche, des distributions vagues (non-informatives) sont utilisées pour les paramètres dont les valeurs sont peu ou mal connues, tels que les coefficients de transfert entre deux compartiments non-physiologiques.

Dans le cas d'études toxicocinétiques, l'information *a priori* disponible concerne la population dans sa globalité, et non les individus eux-mêmes. C'est pourquoi, dans le modèle de population, l'information *a priori* va renseigner les paramètres de population (à savoir, avec les notations proposées,  $\mu$  et  $\Sigma^2$ ).

Le modèle toxicocinétique proposé ci-dessus pour le butadiène ne repose qu'en partie sur des justifications physiologiques, et des distributions uniformes (non-informatives) ont été définies pour les paramètres correspondant aux "moyennes de population". Pour chaque variance de population, la distribution *a priori* choisie était une loi de probabilité gamma inverse, cette loi étant une loi de référence pour les variances inconnues (Bernardo and Smith, 1994).

### *L'estimation des paramètres : les méthodes de Monte Carlo par Chaînes de Markov*

Après la définition des distributions *a priori*, il devient possible de calibrer le modèle à l'aide des données expérimentales. Cependant, les calculs nécessaires par cette calibration sont rarement réalisables analytiquement. Ceci est dû aux formes complexes des distributions *a posteriori* et aux multiples niveaux de hiérarchie du modèle. Pour contourner ces difficultés, les techniques de simulations Monte Carlo par Chaînes de Markov (MCMC) sont utilisées. Les simulations MCMC sont une classe particulière de simulations

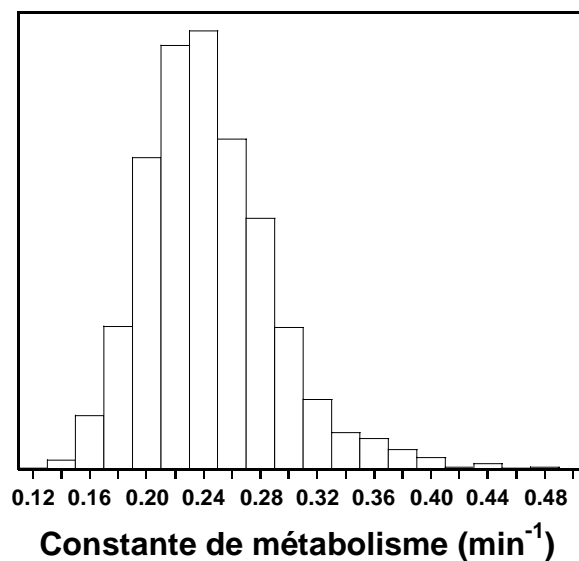
stochastiques. Les simulations Monte Carlo simples sont basées sur des tirages aléatoires successifs et indépendants issus de distributions *a priori*. Les tirages des MCMC sont influencés par les données expérimentales, et le  $n^{\text{ième}}$  tirage est conditionné par le précédent. Les techniques MCMC sont particulièrement intéressantes, dans notre cas, puisqu'elles permettent la simulation d'une loi *a posteriori* sans en connaître la fonction analytique de la densité de probabilité cumulée. Après un certain nombre de tirages, la chaîne simulée converge vers la densité *a posteriori* conjointe (*i.e.* corrélations incluses) des paramètres du modèle. En pratique, plusieurs chaînes sont simulées, permettant de vérifier leur convergence en loi. Différents critères de convergence sont utilisés (Cowles and Carlin, 1996), tels que le critère de Gelman et Rubin (Gelman *et al.*, 1995). En recueillant un échantillon suffisamment grand de valeurs simulées après convergence, des réalisations des distributions *a posteriori* des paramètres sont obtenues.

#### *Analyse des résultats : lois a posteriori*

Une représentation de la distribution conjointe de l'ensemble des paramètres du modèle peut donc être obtenue à partir d'un échantillon de simulations MCMC. Le mode conjoint (vecteur de densité *a posteriori* la plus élevée), est souvent sélectionné pour faire une première vérification du modèle avant d'inférer sur les paramètres. Cette vérification est détaillée dans la section suivante. La distribution marginale (ou individuelle) de chacun des paramètres, correspondant à leur densité de probabilité indépendamment des autres, est également donnée par les simulations MCMC.

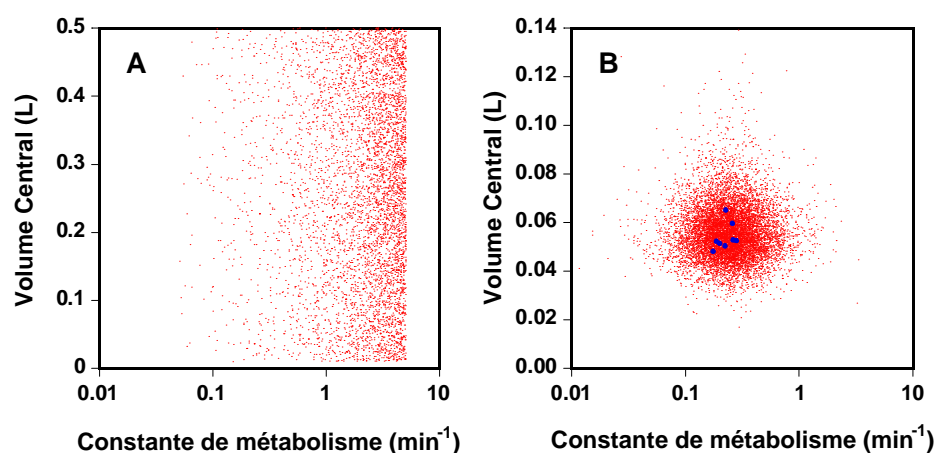
Notre modèle pour le butadiène a été calibré à l'aide des données de concentrations mesurées dans l'air exhalé. Pour chacun des paramètres individuels et de population, une distribution *a posteriori* a été obtenue. La distribution de la moyenne de population de la constante de métabolisme  $k_e$  est présentée Figure 6 sous la forme d'un histogramme créé à partir d'un échantillon de 10000 valeurs générées par MCMC.

Chaque distribution marginale *a posteriori* est généralement moins diffuse que celle définie *a priori*, indiquant que les données expérimentales ont pu informer sur la valeur du paramètre. De ces distributions marginales, les statistiques descriptives (moyenne, l'écart type, quantiles, intervalles de crédibilité sous forme d'intervalles inter-quantiles) de chaque paramètre (individuel ou de population) du modèle peuvent être extraites. Ces distributions peuvent aussi être obtenues pour des prédictions du modèle. Pour obtenir des intervalles de crédibilité sur une valeur prédite (par exemple, la quantité de butadiène métabolisée), il suffit de la calculer pour chaque vecteur de paramètres simulé par MCMC. On obtient un échantillon de valeurs prédites dont on peut déduire l'histogramme, la moyenne, les quantiles, *etc.* Les distributions *a posteriori* sont aussi utilisées lors des étapes ultérieures de vérification, validation ou comparaison de modèles.



**Figure 6 :** *Histogramme de la distribution a posteriori de la moyenne de population de la constante de métabolisme du butadiène.*

La Figure 7 permet de visualiser le processus de calibration bayésienne. Sur la Figure 7A, les distributions *a priori* des paramètres sont uniformes (en échelle logarithmique), alors que sur la Figure 7B les points échantillonnés des distributions *a posteriori* ne sont plus uniformément dispersés, mais concentrés autour d'une valeur moyenne. L'incertitude sur ces paramètres a donc diminué. Par ailleurs, les valeurs échantillonnées, issues de la distribution *a posteriori* des paramètres de population (obtenus par simulations MCMC) sont beaucoup plus dispersées que celles correspondant aux sujets observés. Ce phénomène montre que l'estimation de la variabilité dans la population est plus forte que celle observée dans l'échantillon. Ceci est dû au petit nombre de sujets dans l'échantillon et donc à l'incertitude entachant l'estimation de la moyenne,  $\mu$ , et de la variance,  $\Sigma^2$ , de population.



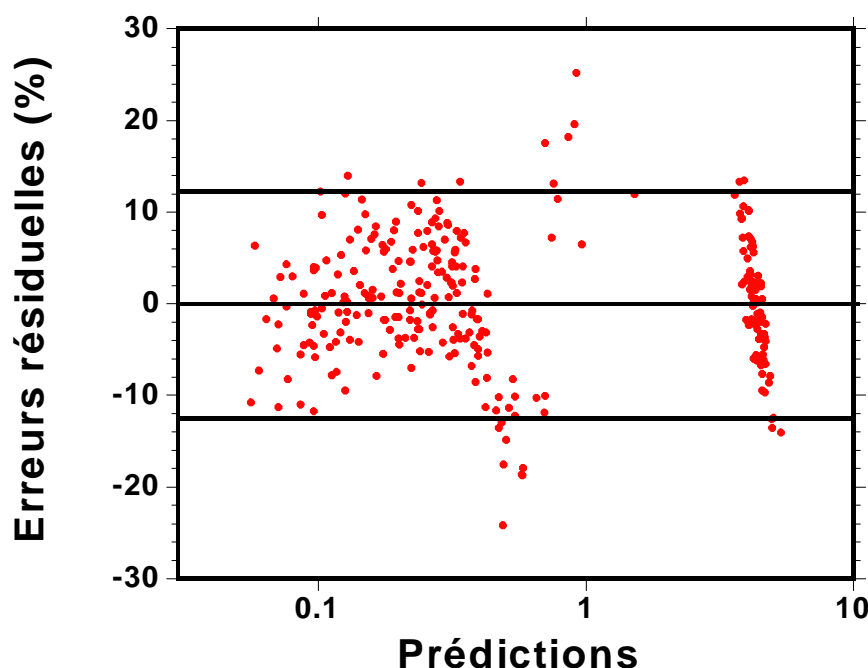
**Figure 7 :** Valeurs de deux paramètres du modèle toxicocinétique (volume du compartiment central, en litre et constante de métabolisme,  $k_e$ , en  $\text{min}^{-1}$ ) avant et après calibration. En A, un échantillon de la distribution *a priori* des paramètres. En B, un échantillon *a posteriori*. Chaque point donne la valeur du volume du compartiment central et la constante de métabolisme pour un sujet simulé, les gros points (en B) correspondent aux sujets observés. Noter le changement d'échelle.

## Vérification du modèle

### Vérification de l'ajustement

Après calibration du modèle, il est nécessaire de vérifier son ajustement aux données. Plusieurs représentations graphiques permettent de visualiser les différences entre les prédictions du modèle et les données expérimentales. La plus répandue superpose les données expérimentales à la courbe d'ajustement du modèle. La Figure 2 présente les mesures expérimentales, ainsi que l'évolution prédite de la concentration de butadiène dans l'air exhalé (l'estimateur pour chaque paramètre est le mode de la densité *a posteriori* conjointe) chez 4 sujets exposés lors de l'étude I. Ensuite, une quantification de l'écart entre les prédictions et les observations peut être évaluée, notamment par le calcul des erreurs résiduelles ou relatives. Pour les expositions au butadiène, les erreurs résiduelles (en terme de pourcentage des valeurs prédites) ont été calculées pour chaque couple (observation, prédiction) de chaque sujet exposé (8 au total) et sont présentées Figure 8. Un intervalle théorique (-12,6%; 12,6%), dans lequel la probabilité d'observer une variable aléatoire normale de moyenne 0 et d'écart-type géométrique égal à 1,077 vaut 90%, a été tracé. Le modèle présuppose que l'erreur de mesure est lognormale et sa calibration a permis d'estimer que son écart-type *a posteriori* vaut en moyenne 1,077. Par conséquent, le modèle peut être considéré comme approximativement vérifié si environ 90% des erreurs résiduelles sont situées dans l'intervalle théorique précédemment défini. Cette vérification n'est qu'approximative car si l'estimation 1.077 était fautive (si le modèle était mauvais), l'intervalle théorique le serait aussi. Ici, 24 erreurs résiduelles sur les 289 totales sont situées hors de cet

intervalle (soit 8,3%) ; le modèle peut donc être considéré comme approximativement correct. On remarque cependant de légères déviations systématiques (de l'ordre de quelques %) que nous avons négligé.



**Figure 8 :** Résidus observés à partir des prédictions du modèle à deux compartiments calibré.

Chaque point représente la différence entre une observation et la prédiction du modèle. Ces différences sont données en terme de pourcentage de prédiction. Les lignes noires représentent l'intervalle théorique à 90% (sous hypothèse de normalité des erreurs de moyenne et d'écart-type respectivement égaux à 0 et 1,077). La valeur de l'écart-type est celle obtenue par calibration du modèle.

#### *Discussion de la valeur des paramètres du modèle*

La seconde étape de la vérification du modèle consiste à discuter les estimations des paramètres en comparant leurs distributions *a priori* et *a posteriori* pour s'assurer de leur cohérence. En effet, si la connaissance *a priori* sur la valeur de certains paramètres est précise (distribution *a priori* informative), la calibration devrait donner une distribution *a posteriori* proche de celle *a priori*. Dans la mesure où ces distributions seraient notablement différentes, il conviendrait de remettre en question le modèle (toxicocinétique et/ou statistique), ou encore les données. Par ailleurs, si la distribution *a priori* est vague, la valeur estimée du paramètre n'apporte que peu d'information pour la vérification du modèle. C'est malheureusement le cas dans l'exemple proposé : Les paramètres du modèle à deux compartiments n'étant pas physiologiques, des distributions *a priori* non informatives ont été choisies.



### ***Validation du modèle***

La finalité de l'expérimentation est souvent la critique de l'adéquation d'une hypothèse ou de résultats théoriques, éventuellement par l'intermédiaire de simulations issues d'un modèle associé. Même si la vérification d'un modèle peut être admise suite à la confrontation des résultats théoriques avec ceux de l'expérimentation, il n'existe pas de méthode universelle permettant de conclure qu'un modèle est valide. La validation est intrinsèquement liée à la critique, et le modèle est généralement critiquable en plusieurs points. Nous allons chercher ici à valider le modèle toxicocinétique. Il aurait aussi été possible de faire porter la critique sur le modèle de population en discutant la nature des distributions statistiques des différents paramètres du modèle.

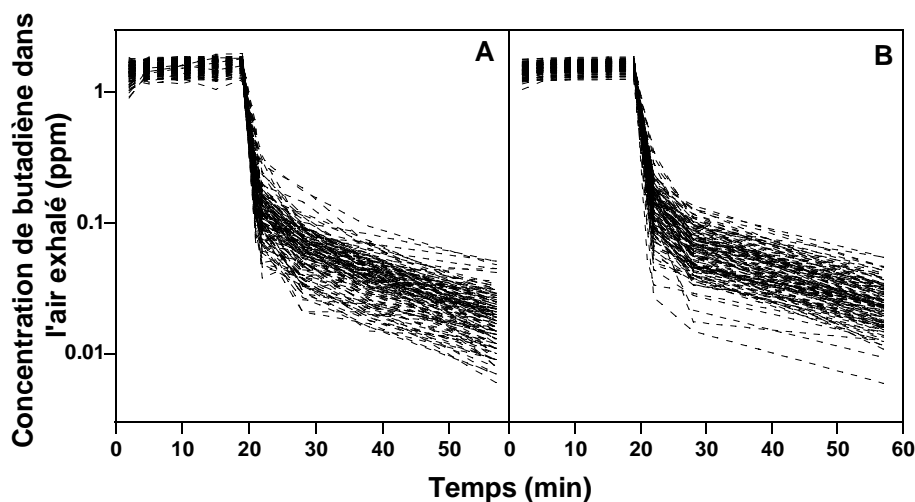
En toxicocinétique, plusieurs propositions concernant la démarche de validation émergent (Chen, 2000; Gargas *et al.*, 2000a, c; Kohn, 1997) et la définition même du terme "validation" anime toujours des débats. Selon nous, ce processus comprend différentes étapes clairement nécessaires même si certaines questions sur l'utilisation des données restent ouvertes :

- Vérifier que la structure du modèle toxicocinétique est en accord avec la connaissance du processus physiologique,
- Définir un ensemble de comportements qualitatifs et quantitatifs à respecter (en dehors de ceux déjà identifiés par les données disponibles),
- Si des données sont disponibles, calibrer et vérifier le modèle
- S'assurer que les prédictions du modèle respectent les comportements requis, en particulier à l'aide de la validation croisée (comparaison des prédictions du modèle avec un jeu de données n'ayant pas été utilisé pour la calibration).

Le processus de validation est donc plus complet que le seul processus de vérification. Il permet de s'assurer que le modèle donne des prédictions correctes en dehors de la plage d'observation des données (mais en restant dans son domaine de validité).

Pour l'étape de validation croisée, dans le cadre de nos études sur le du butadiène, des prédictions de l'évolution de la concentration de butadiène dans l'air exhalé ont été obtenues pour les sujets de l'étude II, simplement sur la base du modèle calibré avec les données de l'étude I. Le comportement de ces courbes est en accord avec notre connaissance du phénomène physiologique (décroissance de la concentration exhalée lors de l'arrêt de l'exposition, par exemple). De plus, les courbes prédites se superposent aux données de l'étude II sans qu'un ajustement soit nécessaire (Figure 9). Ceci constitue un bon argument de validité du modèle.

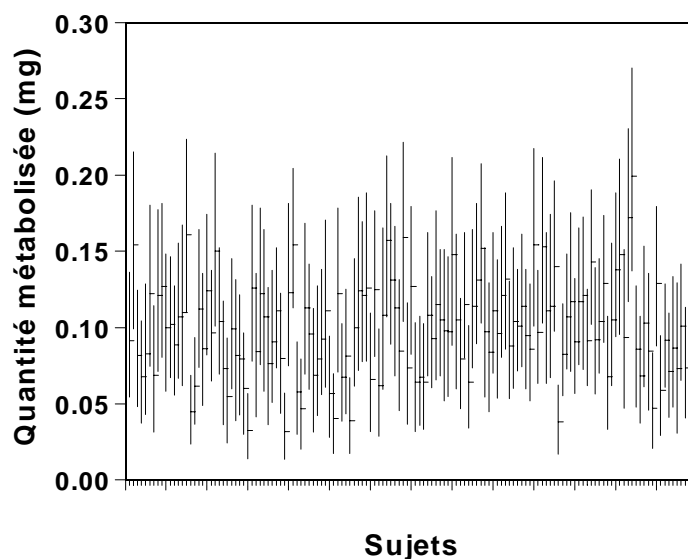
On manque cependant de procédures précises pour l'utilisation des données dans le processus de validation. Il est établi qu'il est préférable de réserver une partie des données pour la validation. Si le modèle est validé, la partie des observations gardée peut être utilisée pour une recalibration finale du modèle. Cette méthodologie mériterait d'être affinée. Par exemple, la fraction de données réservées à la validation ou encore la façon de précise de choisir ces données sont basées sur des règles empiriques, ce qui n'est probablement pas optimal.



**Figure 9 :** Evolution des concentrations de butadiène dans l'air exhalé pendant et après l'exposition de 138 sujets : (A) Les courbes sont les prédictions de concentration exhalées pour les sujets purement simulés à partir du modèle de population (après calibration dans l'étude I), (B) les courbes relient les concentrations de butadiène exhalé mesurées chez les sujets de l'étude II. Chaque courbe correspond à un sujet.

#### *Prédictions à partir d'un modèle validé*

Après validation du modèle, les données relevées sur la totalité des sujets (les 138 de l'étude II) ont été utilisées pour une recalibration du modèle. De ce fait, toute l'information disponible a été utilisée pour renseigner les paramètres du modèle. Partant des simulations MCMC, il est possible d'obtenir des distributions de quantités de butadiène métabolisées par chaque sujet en utilisant les équations du modèle et en affectant aux paramètres les valeurs simulées. Sur la Figure 10, chaque tiret représente la valeur moyenne de la quantité métabolisée et chaque segment l'intervalle de crédibilité à 95% associé. Certains de ces intervalles ne se recouvrent pas, montrant ainsi que les comportements métaboliques diffèrent entre individus.



**Figure 10 :** *Estimation prédictive de la quantité de butadiène métabolisée par chacun des 138 volontaires de l'étude II. Ces prédictions ont été obtenues à partir du modèle toxicocinétique calibré, c'est à dire après estimation des distributions a posteriori de leurs paramètres individuels.*

### **Comparaison de modèles**

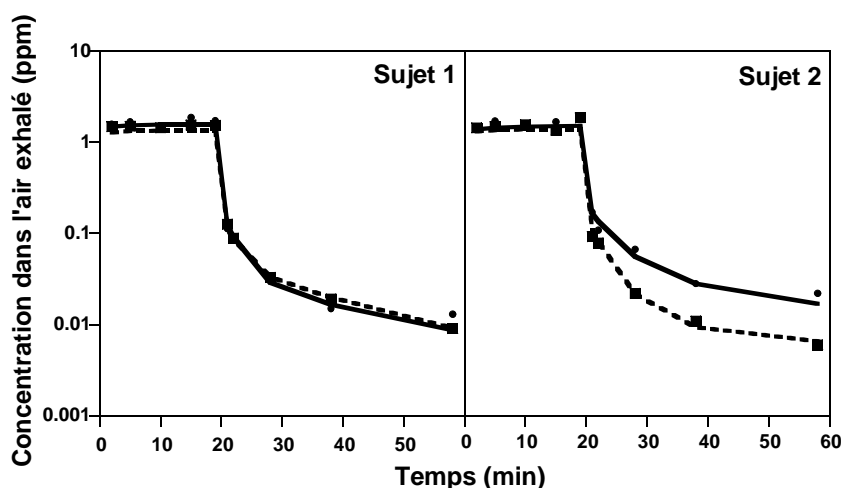
La construction d'un modèle repose principalement sur les hypothèses (mécanistes et statistiques) à partir desquelles il a été fondé. Néanmoins, des ensembles d'hypothèses divergentes peuvent paraître raisonnables, sans qu'aucune raison ne permette d'écarter l'une ou l'autre. Selon les hypothèses retenues, le modèle différera. Lorsque deux modèles sont en compétition, il est possible de les comparer formellement grâce aux facteurs de Bayes afin de choisir en faveur de l'un d'eux, et ainsi d'apprécier leurs hypothèses associées. Les facteurs de Bayes sont le rapport des vraisemblances *a posteriori* de chacun des modèles vérifiés et validés. Une règle de décision simple, reposant sur la valeur de ce rapport, permet ensuite de décider en faveur de l'un des modèles (Bernardo and Smith, 1994). Si le facteur de Bayes d'un modèle A contre un modèle B est supérieur à 1, le modèle A est considéré meilleur et dans le cas contraire, ce sera B. Dans la mesure où plus de deux modèles sont en compétition, la comparaison des facteurs de Bayes permet de la même façon de les comparer (deux à deux par exemple) et de décider en faveur du modèle le plus vraisemblable *a posteriori*.

Nous avons appliqué cette technique pour décider de l'existence ou non d'une évolution dynamique du métabolisme du butadiène pour un même sujet (question toxicologique associée à l'étude III). Cela implique de quantifier la constante de métabolisme lors de différentes expositions d'un même sujet. Onze sujets ont donc été exposés au butadiène à deux occasions différentes, avec recueil des concentrations dans l'air exhalé. Dans un premier temps, le modèle toxicocinétique a été calibré à l'aide des données des deux expositions distinctes de chaque sujet. La Figure 11 présente, pour deux sujets, les prédictions du modèle toxicocinétique

calibré avec chaque ensemble de données. Bien que les deux prédictions pour le premier sujet soient assez semblables (Figure 11, Sujet 1), les prédictions obtenues pour le deuxième sujet sont apparemment différentes (Figure 11, Sujet 2). Deux modèles statistiques ont été proposés pour décrire ces données :

- le modèle A tenant compte uniquement de la variabilité inter-individuelle (l'hypothèse sous-jacente étant l'absence de variabilité intra-individuelle),
- le modèle B prenant en compte la variabilité inter-individuelle et la variabilité intra-individuelle.

Ces deux modèles ont donc été définis dans un contexte de population et induisent des ensembles de distributions *a priori* distincts. La calibration des modèles A et B aux données expérimentales a été menée indépendamment, conduisant à deux ensembles de lois *a posteriori*. Le calcul du facteur de Bayes de B contre A valait 23 et a donc permis de choisir le modèle B incluant les variabilités inter- et intra-individuelle. De ce fait, la présence de différences intra-individuelles significatives dans le métabolisme du butadiène a été mise en évidence.



**Figure 11 :** *Prédictions d'un modèle toxicocinétique pour le butadiène calibré à partir de mesures prélevées chez deux sujets en deux occasions différentes. Les données et les prédictions de l'occasion 1 sont respectivement représentées par des cercles et une courbe pleine. Pour l'occasion 2, elles sont représentées par des carrés et une ligne discontinue.*

### **Optimisation de plans expérimentaux**

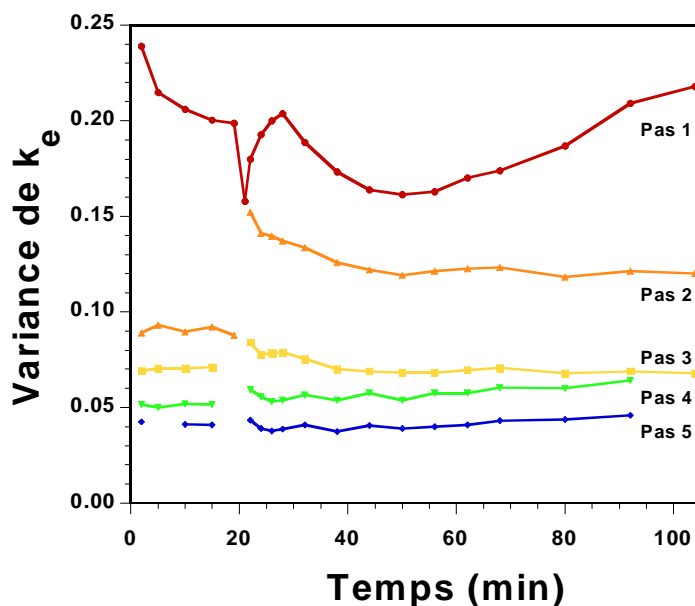
Pour compléter la compréhension d'un phénomène ou l'étudier à plus large échelle, le recueil de données expérimentales supplémentaires à l'étude en cours peut être nécessaire. En général, pour de nouvelles expérimentations, on souhaite optimiser la stratégie de collecte des données de telle sorte qu'à coût donné, elles amènent un maximum d'information. Le "coût" peut inclure les aspects éthiques (risque pour les sujets exposés) ou financiers (nombre d'échantillon à doser, prix du dosage), ou d'autres encore.

Dans le cadre de nos études sur le métabolisme du butadiène, nous avons souhaité estimer la constante de métabolisme  $k_e$  sur un plus grand nombre de sujet que pour l'étude I afin de tester les différences entre sexes et entre ethnies. Pour diminuer le coût de cette large étude et le risque pour ses sujets, tout en obtenant

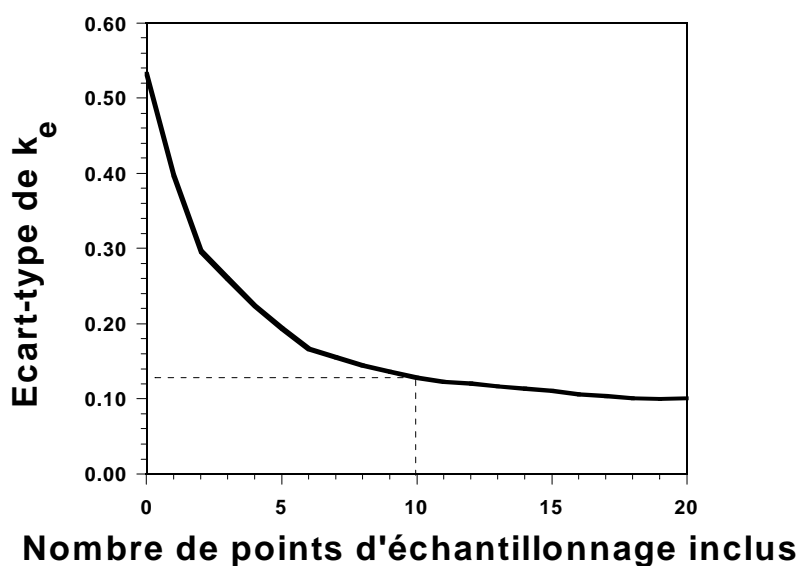
de bonnes estimations des  $k_e$  individuels, un compromis a été cherché entre le nombre de prélèvements à effectuer, leur position dans le temps et la précision de  $k_e$  souhaitée. Cette optimisation du protocole d'échantillonnage repose sur la base de l'information acquise par l'analyse des données expérimentales préliminaires sur les 8 sujets de l'étude I (Bois *et al.*, 1999a).

Pour connaître le nombre de prélèvements à effectuer, la fonction liant la précision (inverse de la variance) de  $k_e$  au nombre de prélèvements d'air exhalé a été déterminée sur la base de simulations (Bois *et al.*, 1999a). Pour ce faire, des données ont été simulées pour les futurs sujets de l'étude, à l'aide du modèle calibré (de façon similaire à ce qui a été présenté dans la section validation, ci-dessus). Les covariables (comme les masses corporelles) de chacun de ces sujets ont été obtenues par tirages aléatoires. Une recalibration du modèle à l'aide des données simulées a donné une nouvelle variance (la moyenne ne changeant pas). Cette estimation a été faite pour différents nombres et temps de mesures par sujet. Pour connaître les temps auxquels les mesures devaient être relevées, l'optimisation a été faite pas à pas sur une grille de 20 points de temps possibles. Chaque pas correspondait à l'introduction d'une nouvelle mesure dont le temps a été choisi comme celui minimisant la variance de  $k_e$  à ce pas. La Figure 12 présente la valeur attendue de la variance de  $k_e$  en fonction des temps auxquels la mesure peut être relevée. A chaque pas, le temps donnant la variance minimum a été inclus dans le plan expérimental. Ainsi un point de mesure à 21 minutes est sélectionné au pas 1, à 19 minutes au pas 2, *etc.* Seuls les 5 premiers pas sont donnés sur la Figure 12. Ceci a permis de déterminer qu'après 10 observations (aux temps optimaux prescrits), pour chaque sujet, l'écart type de  $k_e$  serait de l'ordre de 13% (Figure 13). Cette précision a été jugée satisfaisante (elle est de l'ordre des précisions obtenues dans d'autres études toxicocinétiques). C'est ainsi que l'on a abouti à l'échantillonnage aux temps 2, 5, 10, 15, 19, 21, 22, 28, 38, et 104 minutes pour les sujets de l'étude III.

De nouvelles méthodes "particulières" sont également utilisées pour l'optimisation bayésienne de protocoles expérimentaux (Amzal *et al.*, 2004). Les problèmes d'optimisation peuvent être résolus en explorant une surface d'utilité par simulation MCMC (Mueller, 1999), le plan optimal correspondant au mode de cette surface. Les méthodes particulières permettent d'explorer plus efficacement l'espace de grande dimension des plans expérimentaux.



**Figure 12 :** Estimation de la variance attendue associée à chaque temps de mesure du plan expérimental (depuis le début l'exposition à 2 ppm pendant 20 minutes). Les 5 premiers pas du procédé d'optimisation sont représentés. Les "trous" correspondent aux points inclus au fur et à mesure dans le plan optimal.



**Figure 13 :** Evolution de l'estimation de l'écart-type de  $k_e$  (en échelle log), pour un individu aléatoire, en fonction du nombre de mesures de concentration de butadiène dans l'air exhalé.

## CONCLUSION

La statistique bayésienne est un outil particulièrement adapté à l'analyse des données toxicocinétiques pour diverses raisons. Tout d'abord, elle permet d'incorporer pertinemment une grande partie de l'information disponible dans la littérature. Les valeurs standards définies par différents auteurs ou

commission (par exemple, (International Commission on Radiological Protection (ICRP), 2002)) pour les paramètres permettent de définir des distributions *a priori* informatives décrivant l'incertitude sur ces grandeurs. De ce fait, l'estimation bayésienne des paramètres limite le besoin de recourir à l'expérimentation en utilisant l'information disponible *a priori* sur la valeur des paramètres physiologiques des modèles toxicocinétiques (flux sanguin, volume des organes du corps, *etc.*). Elle répond ainsi au problème récurrent de la pauvreté des données imposée par les contraintes éthiques ou celles de coût.

Cette approche, facilement réalisable grâce à certains logiciels (Bois and Maszle, 1997a; Thomas *et al.*, 1992), permet également d'estimer les paramètres de modèles complexes tels que les modèles hiérarchiques de population. La construction d'un modèle de population dans le cadre de la statistique classique se heurte souvent à la difficulté d'estimer ses paramètres. Dans ce contexte, des techniques d'estimation approximative, comme la méthode NONMEM (Graves and Chang, 1988; Kaniwa *et al.*, 1990) ou NMPC (Sheiner *et al.*), ont été proposées et peuvent être mises en place en routine grâce à certains logiciels. L'approche bayésienne permet de se passer de ces approximations et offre une méthode d'inférence assez simple : Une fois les distributions *a priori* définies, la calibration fournit ainsi une estimation des paramètres individuels et de population sous forme de distributions de probabilités, desquelles peuvent être déduits des intervalles de crédibilité. Ces estimations sont obtenues en tenant compte des divers niveaux d'incertitude et de variabilité inhérentes aux données toxicocinétiques.

L'ensemble de ces travaux peut être étendu dans différentes directions :

- coupler la modélisation toxicocinétique et l'imagerie médicale en travaillant sur des données de distribution des substances dans le corps entier. Dans ce cadre, la statistique bayésienne permettrait de calibrer des modèles réalistes, spécifiques d'individus précis. Tirer parti des avantages des méthodes de traitement bayésiennes de données spatiales est un axe de recherche actuel.
- La modélisation pharmacocinétique s'oriente vers des descriptions de plus en plus fines de l'organisme, capables d'intégrer davantage de données. La création de modèles toxicocinétiques à la fois génériques et détaillés permettrait de décrire, à l'aide d'un même modèle, la distribution de multiples substances et leurs interactions.
- Bien que l'information sur les valeurs moyennes (et éventuellement les variances) des paramètres physiologiques dans la population sont, la plupart du temps, disponibles dans la littérature, les informations sur la nature de leurs distributions sont encore assez pauvres. La forme précise des lois de distributions est quasiment toujours inconnue. Les modèles de population généralement utilisés pour l'analyse des données toxicocinétiques supposent connues les lois de distribution des paramètres de population. En général, des lois normales, log-normales ou Gamma sont utilisées. Ces hypothèses mériteraient d'être relaxées par l'utilisation de modèles de population non-paramétrique (Claret and Iliadis, 1996; Wakefield and Bennett, 1996).
- Finalement, une réflexion sur la mesure de l'adéquation des modèles aux données qu'ils ont pour objectif de décrire ou de prédire devrait être menée. La démarche proposée dans cet article ne couvre qu'en partie le processus de validation. La validation d'un modèle dans le contexte bayésien nécessite encore des

réflexions de fond sur le processus lui-même. Enfin, le terme même de validation peut sous-entendre un caractère absolu sur stabilité, dans le temps, des phénomènes étudiés, ce constitue une hypothèse forte. Prendre en compte l'aspect dynamique de la réalité est pourtant envisageable. Par exemple, la distribution d'un paramètre de population (la masse corporelle moyenne, par exemple) peut évoluer au cours du temps pendant le recueil (plus ou moins long) des données. Pour tenir compte de ces évolutions, de nouveaux algorithmes stochastiques (Tierney and Mira, 1999) (les méthodes "particulaires") permettent de prendre en compte l'évolution à long terme des paramètres, dans le cas d'analyses séquentielles des données.

## ANNEXE : EQUATIONS DEFINISSANT LE MODELE A 2-COMPATIMENTS.

La quantité de 1,3-butadiène dans les compartiments central ( $Q_{cent}$  en mmoles) et périphérique ( $Q_{periph}$  en mmoles), et la quantité métabolisée ( $Q_{met}$  en mmoles) sont données par :

$$\begin{aligned}\frac{\partial Q_{cent}}{\partial t} &= K_{in} C_{inh} + \frac{K_{cp}}{P_{pc}} \frac{Q_{periph}}{V_p} - \left( K_{cp} + \frac{K_{in}}{P_{ca}} \right) \frac{Q_{cent}}{V_c} - K_{met} Q_{cent} \\ \frac{\partial Q_{periph}}{\partial t} &= K_{cp} \frac{Q_{cent}}{V_c} - \frac{K_{cp}}{P_{pc}} \frac{Q_{periph}}{V_p} \\ \frac{\partial Q_{met}}{\partial t} &= K_{met} Q_{cent}\end{aligned}$$

où  $C_{inh}$  est la concentration inhalée (en mmoles/L);  $K_{in}$  est le taux de ventilation alvéolaire par minute (en L/min);  $K_{cp}$  est la constante de transfert du compartiment central vers le compartiment périphérique (en L/min);  $P_{pc}$  est le coefficient de partage du tissu périphérique sur le tissu central (sans dimension);  $V_c$  et  $V_p$  sont respectivement les volumes des compartiments central et périphérique (en L);  $P_{ca}$  est le coefficient de partage tissu central sur air (sans dimension);  $K_{met}$  est la constante de métabolisme (en 1/min). La concentration de butadiène dans l'air exhalé (en mmoles/L) est donnée par :

$$C_{exh} = \frac{0.7 \times Q_{cent}}{P_{ca} \times V_c} + 0.3 \times C_{inh}$$

Ce qui suppose un espace mort pulmonaire de 30%.

Le volume du compartiment central a été mis à l'échelle de la masse corporelle,  $MC$  (en kg), et son facteur d'échelle,  $sc\_V_c$ , était le paramètre estimé:

$$V_c = sc\_V_c \times MC$$

## REFERENCES

- Aarons, L. (1992). Population pharmacokinetics. *International Journal of Clinical Pharmacology, Therapy, and Toxicology* 30, 520-522.
- Amzal, B., Bois, F. Y., Parent, E., and Robert, C. P. (2004). Bayesian Optimal Design via Interacting MCMC. *Cahiers du Ceremade* 034.



- Beal, S. L., and Sheiner, L. B. (1982). Estimating population kinetics. *CRC Critical Reviews in Biomedical Engineering* 8, 195-222.
- Bernardo, J. M., and Smith, A. F. M. (1994). *Bayesian Theory*. Wiley, New York.
- Bernillon, P., and Bois, F. Y. (2000). Statistical issues in toxicokinetic modeling: A Bayesian perspective. *Environmental Health Perspectives* 108, 883-893.
- Bois, F. Y. (2003). Modélisation toxicocinétique de la concentration sanguine de 2,3,7,8-tetrachloro-p-dioxine après ingestion chez la femme. *Environnement, Risques et Santé* 2.
- Bois, F. Y., Gelman, A., Jiang, J., Maszle, D. R., Zeise, L., and Alexeef, G. (1996). Population toxicokinetics of tetrachloroethylene. *Archives of Toxicology* 70, 347-355.
- Bois, F. Y., and Maszle, D. (1997). MCSim: a simulation program. *Journal of Statistical Software* 2(9), [http://toxi.ineris.fr/activites/toxicologie\\_quantitative/mcsim/mcsim.php](http://toxi.ineris.fr/activites/toxicologie_quantitative/mcsim/mcsim.php).
- Bois, F. Y., Smith, T. J., Gelman, A., Chang, H. Y., and Smith, A. E. (1999). Optimal design for a study of butadiene toxicokinetics in humans. *Toxicological Sciences: an Official Journal of the Society of Toxicology* 49, 213-224.
- Brown, R. P., Delp, M. D., Lindstedt, S. L., Rhomberg, L. R., and Beliles, R. P. (1997). Physiological parameter values for physiologically based pharmacokinetic models. *Toxicology and Industrial Health* 13, 407-484.
- Chen, C. (2000). Validation of a population pharmacokinetic model for adjunctive lamotrigine therapy in children. *British Journal of Clinical Pharmacology* 50, 135-145.
- Claret, L., and Iliadis, A. (1996). Nonparametric density estimation applied to population pharmacokinetics. *Math Biosci* 133, 51-68.
- Cowles, M. K., and Carlin, B. P. (1996). Markov chain Monte Carlo convergence diagnostics: a comparative review. *Journal of the American Statistical Association* 91, 883-904.
- Gargas, M. L., Tyler, T. R., Sweeney, L. M., Corley, R. A., Weitz, K. K., Mast, T. J., Paustenbach, D. J., and Hays, S. M. (2000a). A toxicokinetic study of inhaled ethylene glycol ethyl ether acetate and validation of a physiologically based pharmacokinetic model for rat and human. *Toxicol Appl Pharmacol* 165, 63-73.
- Gargas, M. L., Tyler, T. R., Sweeney, L. M., Corley, R. A., Weitz, K. K., Mast, T. J., Paustenbach, D. J., and Hays, S. M. (2000b). A toxicokinetic study of inhaled ethylene glycol monomethyl ether (2-ME) and validation of a physiologically based pharmacokinetic model for the pregnant rat and human. *Toxicol Appl Pharmacol* 165, 53-62.
- Gelman, A., Bois, F. Y., and Jiang, J. (1996). Physiological pharmacokinetic analysis using population modeling and informative prior distributions. *Journal of the American Statistical Association* 91, 1400-1412.
- Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (1995). *Bayesian Data Analysis*. Chapman & Hall, London.
- Gibaldi, M., and Perrier, D. (1982). *Pharmacokinetics*. Marcel Dekker, New York.

Graves, D. A., and Chang, I. (1988). Application of NONMEM to routine bioavailability data. *Journal of Pharmacokinetics and Biopharmaceutics* 18, 145-160.

International Agency for Research on Cancer (IARC) (1992). Occupational Exposures to Mists and Vapours from Strong Inorganic Acids and Other Industrial Chemicals. International Agency for Research on Cancer, Lyon, France.

International Commission on Radiological Protection (ICRP) (2002). Basic Anatomical and Physiological data for Use in Radiological Protection: Reference Values.

Kaniwa, N., Aoyagi, N., Ogata, H., and Ishii, M. (1990). Application of the NONMEM method to evaluation of the bioavailability of drug products. *Journal of Pharmaceutical Sciences* 79, 1116-1120.

Kohn, M. C. (1997). The importance of anatomical realism for validation of physiological models of disposition of inhaled toxicants. *Toxicology and Applied Pharmacology* 147, 448-458.

Ludemann, L., Hamm, B., and Zimmer, C. (2000). Pharmacokinetic analysis of glioma compartments with dynamic Gd-DTPA-enhanced magnetic resonance imaging. *Magn Reson Imaging* 18, 1201-14.

Mueller, P. (1999). Simulation Based Optimal Design. In *Bayesian Statistics 6* (J. M. B. J.O. Berger, A.P. Dawid and A.F.M. Smith, ed., pp. 459-474. Oxford University Press.

Rowland, M., Benet, L. Z., and Graham, G. G. (1973). Clearance concepts in pharmacokinetics. *Journal of Pharmacokinetics and Biopharmaceutics* 1, 123-136.

Sheiner, L. B., Beal, S., and Weichman, B. UCSF NONMEM software. licensed and distributed by GloboMax LLC, <http://c255.ucsf.edu/nonmem0.html>.

Sheiner, L. B., and Beal, S. L. (1980). Evaluation of methods for estimating population pharmacokinetic parameters - I. Michaelis-Menten model: routine clinical pharmacokinetic data. *Journal of Pharmacokinetics and Biopharmaceutics* 8, 553-571.

Thomas, A., Spiegelhalter, D. J., and Gilks, W. R. (1992). BUGS: a program to perform Bayesian inference using Gibbs sampling. In *Bayesian Statistics 4* (A. F. M. Smith, ed., pp. 837-842. Oxford University Press, Oxford.

Tierney, L., and Mira, A. (1999). Some adaptive Monte Carlo methods for Bayesian inference. *Statistics in Medicine* 18, 2507-2515.

Tozer, T. N. (1981). Concepts basic to pharmacokinetics. *Pharmacology and Therapeutics* 12, 109-131.

Wakefield, J., and Bennett, J. (1996). The Bayesian modeling of covariates for population pharmacokinetic models. *Journal of the American Statistical Association* 91, 917-927.

## **5. Modélisation physiologique de données spatio-temporelles d'imagerie médicale**

Nous présentons, ici, le développement de notre premier protocole expérimental pour la paramétrisation des modèles physiologiques toxico/pharmacocinétique : l'intégration de données d'imagerie médicale. Dans le cadre de ce travail, nous avons collaboré avec l'équipe "Biodistribution de marqueurs, applications aux tumeurs solides" (unité INSERM U494). Cette équipe nous a fourni des images RMN de biodistribution d'un agent de contraste (le Vistarem<sup>®</sup>) chez des souris sur lesquelles une tumeur a été greffée. Nous avons ensuite développé un modèle physiologique pharmacocinétique pour cet agent de contraste, permettant ainsi de caractériser la vascularisation de la tumeur. L'article scientifique (soumis à *Magnetic Resonance in Medicine*) est présenté, ainsi qu'un résumé en français.

## **Bayesian physiological pharmacokinetic modeling of tumors' microcirculation magnetic resonance imaging data: Application to the evaluation of an antiangiogenic treatment.**

Céline Brochot<sup>1</sup>, Bertrand Bessoud<sup>2</sup>, Daniel Balvay<sup>2</sup>, Charles-André Cuénod<sup>2,3</sup>, Nathalie Siauve<sup>2,3</sup> and Frédéric Y. Bois<sup>1</sup>

<sup>1</sup>INERIS, Institut National de l'Environnement Industriel et des Risques, Unité de Toxicologie Expérimentale, Parc ALATA BP 2, 60550 Verneuil en Halatte, France

<sup>2</sup>Laboratoire de Recherche en Imagerie, LRI-U494, Faculté Necker, Paris, France

<sup>3</sup>Service de Radiologie, Hôpital Européen Georges Pompidou, 20 rue Leblanc, 75015 Paris, France

Correspondence: Frédéric Y. Bois, Institut National de l'Environnement Industriel et des Risques, Unité de Toxicologie Expérimentale, Parc Alata BP2, 60550 Verneuil en Halatte, France

Phone : + 33 3 44 55 65 96

Fax : + 33 3 44 55 66 05

Soumis à *Magnetic Resonance in Medicine*

## RESUME

Un modèle physiologique pharmacocinétique a été défini pour caractériser et quantifier la microcirculation tumorale chez des souris "nude" ayant une tumeur greffée. Les cinétiques d'un agent de contraste (Vistarem<sup>®</sup>), appartenant à la classe d'agents avec une clairance rapide, ont été mesurées par imagerie par résonance magnétique. Les augmentations du signal ont été enregistrés dans le sang artériel, et dans les tissus de la tumeur pendant 580 secondes après le bolus de Vistarem<sup>®</sup>. Notre approche a permis de construire un modèle "corps entier" de la biodistribution de cet agent, pour lequel les paramètres ont une signification physiologique. Le modèle comprend six compartiments: le plasma artériel et veineux, la tumeur (divisée en capillaires et interstitium), et le reste du corps (divisé aussi en capillaires et interstitium).

Comme application, les paramètres de microcirculation tumorale ont été évalués chez des souris traitées avec un médicament anti-angiogénique (ZD4190), ou un placebo. Trois paramètres de microcirculation tumorale sont étudiés : le flux sanguin, la perméabilité des capillaires, et la fraction volumique de sang. L'analyse des données a été réalisée dans un contexte bayésien, et le modèle a été ajusté aux données expérimentales avec les techniques de Monte Carlo par Chaînes de Markov.

Les résultats montrent que les paramètres de microcirculation tumorale des souris traitées sont globalement plus faibles que ceux des souris non-traitées. Pour le flux sanguin, seules 18% des estimations des souris traitées sont plus grandes que celles des souris non-traitées. Ce pourcentage tombe à 6% pour la perméabilité et la fraction volumique des capillaires.

## ABSTRACT

A physiological pharmacokinetic model was used to estimate tumor microcirculation in nude mice with a grafted tumor. Kinetics of a rapid clearance blood pool agent, Vistarem<sup>®</sup>, were investigated by dynamic MRI. Signal enhancements were recorded in arterial blood and in tumor tissue during 580 seconds after the Vistarem<sup>®</sup> bolus. This approach has allowed whole-body modeling of the agent's biodistribution and a physiological definition of the model's parameters. The model included six compartments: arterial and venous plasma, tumor (split into capillaries and interstitium) and the rest of the body (also split into capillaries and interstitium).

As an application, changes in tumor microcirculation parameters were evaluated in mice receiving either an antiangiogenic treatment (ZD4190) or a placebo. The tumor microcirculation parameters were blood flow, capillary permeability and fractional blood volume. The analysis was performed in a Bayesian framework, and the model was fitted to experimental data using Markov Chain Monte Carlo techniques.

Results showed that tumor microcirculation parameters of the treated mice were globally lower than those in untreated mice. For blood flow, only 18% of the estimates of treated mice were higher than those for untreated mice. This percentage fell to 6% for capillary permeability and plasmatic volume fraction.

**Keywords:** tumor microcirculation; Bayesian analysis; PBPK model; Vistarem<sup>®</sup>; dynamic MRI.

## INTRODUCTION

Dynamic contrast-enhanced magnetic resonance imaging (MRI) has important applications in characterizing tumors' microcirculation with MRI diffusable agents (such as gadopentetate dimeglumine, Gd-DTPA) (Daldrup *et al.*, 1998; Taylor *et al.*, 1999; Tofts *et al.*, 1995). Three quantities are usually used to characterize tumor vasculature: blood flow, vascular volume and vascular permeability (Gillies *et al.*, 2000; Pradel *et al.*, 2003). Typically, a series of measurements is performed at least in the tumor, for each experimental subject. These measurements are then related to the quantities of interest through compartment models, describing the exchanges of the contrast agent between a central compartment (usually plasma) and a peripheral compartment (extracellular space or interstitial water) (Tofts, 1997). In early pharmacokinetic analyses of MRI data (Brix *et al.*, 1991; Tofts and Kermode, 1991), measurements were only performed in tumors. Plasma concentrations were inferred rather than observed. Recent work (Port *et al.*, 2001c; Port *et al.*, 1999) has emphasized the need to monitor arterial blood concentration (or whatever input function) for each individual if accurate estimates of individual microcirculation parameters are required.

We propose here to jointly describe, through an original physiologically based pharmacokinetic (PBPK) model, the time course of a MRI contrast agent, Vistarem<sup>®</sup> (Port *et al.*, 2001a; Port *et al.*, 2001b), in the whole body and a tumor. Our approach offers: a global physiological description of the distribution and elimination of the contrast agent, an integrated analysis of all the data, and a meaningful interpretation of the parameters (*e.g.*, blood flows). So far, the limited use of the PBPK modeling in pharmacokinetic studies was due to the risk of developing overparameterized models (Ludden *et al.*, 1995) for which conventional estimation procedures fail. We avoid this by conducting our analysis within a Bayesian framework (Gelman *et al.*, 1996b), allowing us to include *a priori* information about physiological parameter values. A measurement error model was also defined to take into account errors in the imaging sequence recording.

As an application, we compare tumor microcirculation reduction induced in mice by an antiangiogenic drug (ZD4190, AstraZeneca Pharmaceuticals) on reducing tumor's microcirculation in mice (Pradel *et al.*, 2003; Wedge *et al.*, 2000) to the effects of a placebo.

## METHODS

### *Experimental data*

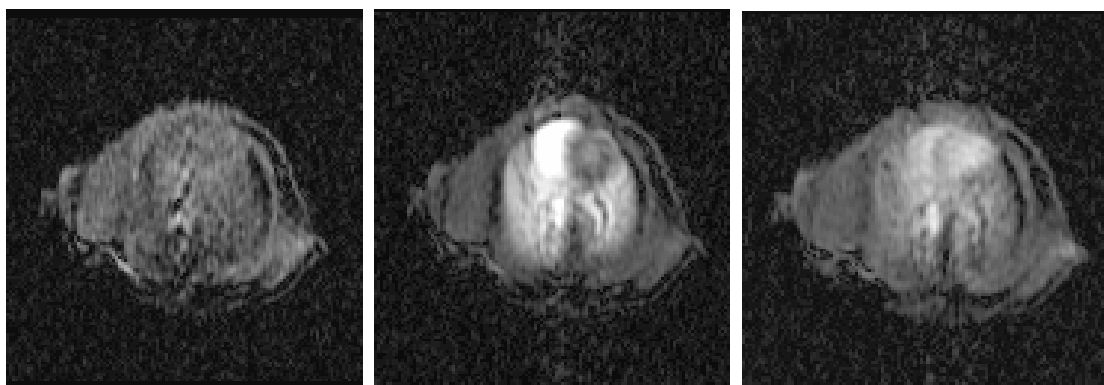
#### *Animals*

Experiments were performed on 13 nude nu/nu male mice (Laboratoire Iffa Credo, L'Arbresle, France), in full compliance with the National Institutes of Health recommendations for animal care. Approximately  $1.5 \times 10^6$  PC-3 human tumor cells were implanted subcutaneously into the flank of each mouse, as described in (Pradel *et al.*, 2003). Tumors were allowed to grow until they reached a diameter of at least 8 mm. Mice were then randomized in two groups: 7 mice treated (T) with VEGFR-2 receptors' inhibitor ZD4190 dissolved in water with 1% polysorbate (AstraZeneca Pharmaceuticals, Macclesfield,

Cheshire, UK), and 6 mice receiving a 1% polysorbate placebo (P). Both were administered orally twice to each mouse, 24 hours and 2 hours before the MRI examination. On the day of examination, mice of Group T weighted 29.6 g (SD 1.5) and were 68- to 84-day-old, mice of Group P weighted 29.6 g (SD 2.6) and were 70- to 85-day-old. For imaging, the animals were anesthetized by an peritoneal injection of ketamine (Rompun<sup>®</sup>, Bayer AG, Leverkusen, Germany) and xylazine (Imalgène<sup>®</sup>, Merial, Lyon, France).

#### *MRI examination*

MRI examination was performed using a 1.5 T system (Sigma, General Electrics, Milwaukee, USA) and a custom small animal dedicated coil. A sagittal 2D T1-weighted spin echo sequence (TE 11ms, TR 400ms, FOV 8x8 cm, 256x128 matrix, 1 NEX) was used to check the adequate positioning of the animal and to select the axial plan level containing the left ventricle cavity and one or two flank tumors. The dynamic acquisition was performed using a single slice T1-weighted 2D Fast Spoiled Gradient-Recalled (FSPGR) sequence: TR 15 ms, TE 2.2 ms, flip angle  $\alpha$  60°, bandwidth 31.25 kHz, 256x128 matrix, FOV 7x3 cm, 5 mm slice). The single slice was positioned at the level selected by the previous sagittal sequence and dynamic acquisition was performed with 10 images baseline and after a caudal vein bolus injection of 0.045 mmolGd/kg of a macromolecular contrast agent (Vistarem<sup>®</sup>, Guerbet, Aulnay-Sous-Bois, France). A series of 3 MR images is shown in Figure 1.



**Figure 1:** *A series of three MR images of Vistarem<sup>®</sup> distribution in a mouse with a grafted tumor: before bolus (baseline), during bolus, and 250 seconds after bolus.*

#### *MRI data*

Two regions of interest (ROIs) were defined on the MR images: one in the left ventricle cavity and the other in the tumor. In each ROI, average signal intensities were measured. After sequence recording of each animal, phantom experiments with 0.112–2.24 mM Vistarem<sup>®</sup> solutions were performed to calibrate a relationship between signal intensities (SI) and relaxation rates. Concentrations,  $C(t)$ , were related to relaxation rates,  $1/T_1$ , according to:



$$\frac{1}{T_1} = \frac{1}{T_0} + r_1 \times C(t) \quad (1)$$

where  $1/T_0$  is the baseline relaxation rate and  $r_1$  is the Vistarem<sup>®</sup> relaxivity at 1.5 Tesla. For each animal, a set of signal intensities and their corresponding relaxation rates were then available. Assuming that T2\* effects are negligible, the inverse equation of the 2D FSPGR sequence signal:

$$\frac{1}{T_1} = \frac{1}{TR} \ln \left( \frac{1 - SI \times \cos \alpha}{1 - SI} \right) \quad (2)$$

was fitted to this dataset. Thanks to this fitted equation, signal intensities recorded in ROIs were converted in relaxation rates (the experimental data in the following).

## Models

### The pharmacokinetic model

Vistarem<sup>®</sup> is a rapid clearance blood pool agent (RCBPA) (Port *et al.*, 2001b). It has limited diffusion across normal endothelium (due to its large molecular size), it is mostly eliminated in urine with clearance equivalent to the glomerular filtration rate. Considering these properties, the four compartments PBPK model developed includes: arterial and venous plasma, tumor and rest of the body (Figure 2). The latter two were split into two subcompartments: capillaries and interstitium (there is no appreciable diffusion of Vistarem<sup>®</sup> in the intra-cellular space).

The Vistarem<sup>®</sup> was introduced in the mouse by bolus injection in the caudal vein. However, if the injection is modeled in the venous compartment, Vistarem<sup>®</sup> kinetics cannot be well reproduced (results not shown). Injection in the caudal vein is physiologically equivalent to a delayed injection shortly before the left ventricle and is well approximated by an infusion in arterial blood. The model leads to the following set of differential equations:

$$\left\{ \begin{array}{l} \frac{dC_{AP}}{dt} = \frac{K_P}{V_{AP}} \times (C_{VP} - C_{AP}) + \frac{K_{Inf}}{V_{AP}} \\ \frac{dC_{BC}}{dt} = \frac{K_{BC}}{V_{BC}} \times (C_{AP} - C_{BC}) + \frac{K_{BCI}}{V_{BC}} \times (C_{BI} - C_{BC}) - \frac{K_E}{V_{BC}} \times C_{BC} \\ \frac{dC_{BI}}{dt} = \frac{K_{BCI}}{V_{BI}} \times (C_{BC} - C_{BI}) \\ \frac{dC_{TC}}{dt} = \frac{K_{TC}}{V_{TC}} \times (C_{AP} - C_{TC}) + \frac{K_{TCI}}{V_{TC}} \times (C_{TI} - C_{TC}) \\ \frac{dC_{TI}}{dt} = \frac{K_{TCI}}{V_{TI}} \times (C_{TC} - C_{TI}) \\ \frac{dC_{VP}}{dt} = \frac{K_{BC}}{V_{VP}} \times C_{BC} + \frac{K_{TC}}{V_{VP}} \times C_{TC} - \frac{K_P}{V_{VP}} \times C_{VP} \end{array} \right. \quad (3)$$

where  $C$  stands for Vistarem<sup>®</sup> concentrations and  $V$  for tissue volumes. Indices  $AP$ ,  $BC$ ,  $BI$ ,  $TC$ ,  $TI$  and  $VP$  correspond respectively to arterial plasma, body's capillaries plasma, body's interstitium, tumor capillaries

plasma, tumor interstitium and venous plasma. Plasma flows are:  $K_P$  for total plasma,  $K_{BC}$  for body plasma, and  $K_{TC}$  for tumor plasma. Diffusion rate constants from capillaries to tumor interstitium and from capillaries to interstitium in the rest of the body are respectively  $K_{TCI}$  and  $K_{BCI}$ .  $K_E$  is the renal elimination rate constant.  $K_{Inf}$  is the infusion rate of the agent in arterial plasma, defined as follow:

$$K_{Inf} = \frac{Q_{Inj}}{T_{EndInf} - T_{BegInf}} \quad (4)$$

where  $Q_{Inj}$  is the quantity injected, and  $T_{BegInf}$  and  $T_{EndInf}$  the beginning and the end of infusion in the arterial plasma. Concentrations are given in mg/ml, volumes in ml, plasma flows in ml/s, transfer constant rates in ml/s, the constant of renal elimination in ml/s, the injection rate in mg/s and the quantity injected in mg.

Preliminary work (not reported) had shown strong correlations between  $K_{TC}$  or  $K_{TCI}$  and  $V_T$ . To deconstrain the estimation of these parameters, two new parameters,  $K_{TC/T}$  and  $K_{TCI/T}$ , were defined:

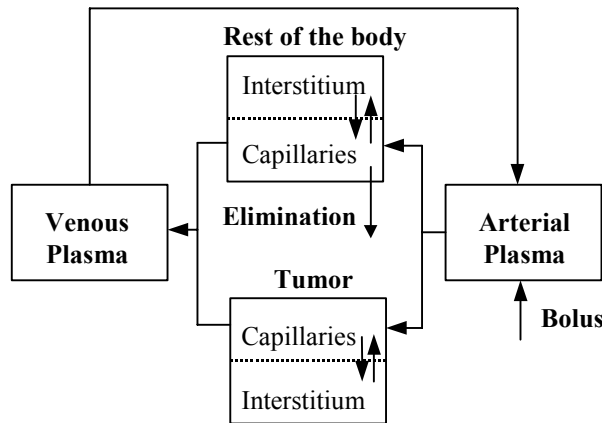
$$K_{TC/T} = \frac{K_{TC}}{V_T} \quad (5)$$

$$K_{TCI/T} = \frac{K_{TCI}}{V_T} \quad (6)$$

and the forth and fifth equations of the system (3) become:

$$\frac{dC_{TC}}{dt} = \frac{K_{TC/T} \times V_T}{V_{TC}} \times (C_{AP} - C_{TC}) + \frac{K_{TCI/T} \times V_T}{V_{TC}} \times (C_{TI} - C_{TC}) \quad (7)$$

$$\frac{dC_{TI}}{dt} = \frac{K_{TCI/T} \times V_T}{V_{TI}} \times (C_{TC} - C_{TI}) \quad (8)$$



**Figure 2:** Physiologically pharmacokinetic model used to describe the distribution of contrast agent Vistarem®. Six compartments are described: arterial and venous plasma, tumor (split in capillaries and interstitium) and rest of the body (also split in capillaries and interstitium). Arrows represent exchanges of material in and out of compartments.

*Relaxation rates - Concentration model*

To fit the pharmacokinetic model, it is necessary to link relaxation rates to concentrations. Eq. [1] can be rewritten as:

$$\Delta R_1 = r_1 \times C(t) \quad (9)$$

with

$$\Delta R_1 = \frac{1}{T_1} - \frac{1}{T_0} \quad (10)$$

On MR images, no distinction can be made at the pixel level between blood, plasma, cells, etc. Since the distribution of Vistarem<sup>®</sup> is extra-cellular, its quantity in blood cells is null. Therefore the relationship between the arterial blood data,  $(\Delta R_1)_{Art}$ , and the arterial plasma concentration,  $C_{AP}$ , is:

$$(\Delta R_1)_{Art} = r_1 \times (1 - \text{Hct}) \times C_{AP}(t) \quad (11)$$

where Hct is the hematocrit, equal to 0.45 in mice (Davis and Morris, 1993).

We also assume that tumor relaxation rates,  $(\Delta R_1)_{Tum}$ , are directly related to average tumor concentrations,  $C_T$ :

$$(\Delta R_1)_{Tum} = r_1 \times C_T(t) \quad (12)$$

with

$$C_T = \frac{Q_{TC} + Q_{TI} + Q_{TCe}}{V_T} \quad (13)$$

where  $V_T$  is the tumor volume,  $Q_{TC}$ ,  $Q_{TI}$  and  $Q_{TCe}$  are respectively Vistarem<sup>®</sup> quantities in the tumor capillaries, interstitium and cellular space. Since  $Q_{TCe}$  is null, we have:

$$(\Delta R_1)_{Tum} = \frac{r_1}{V_T} \times (C_{TC} \times V_{TC} + C_{TI} \times V_{TI}) \quad (14)$$

*Parameterization*

To take into account *a priori* information, some model parameters were scaled. Blood volume ( $V_B$ ) was scaled to bodyweight,  $Bw$ , and its scaling coefficient  $sc\_V_B$  was the parameter actually estimated:

$$V_B = sc\_V_B \times Bw \quad (15)$$

Plasma volume was obtained as:

$$V_P = V_B \times (1 - \text{Hct}) \quad (16)$$

Total body plasma includes arterial, venous and capillary plasma. We scaled  $V_{AP}$  to  $V_P$ :

$$V_{AP} = sc\_V_{AP} \times V_P \quad (17)$$

Similarly,  $V_{BC}$  was scaled to total plasma volume:

$$V_{BC} = sc\_V_{BC} \times V_P \quad (18)$$

and  $V_{TC}$  to tumor volume ( $V_T$ ):

$$V_{TC} = sc\_V_{TC} \times V_T \quad (19)$$

The volume of venous plasma ( $V_{VP}$ ) was obtained by

$$V_{VP} = V_P - V_{AP} - V_{BC} - V_{TC} \quad (20)$$

$V_{BI}$  was scaled to bodyweight with parameter  $sc\_V_{BI}$  and  $V_{TI}$  was scaled to the tumor volume with parameter  $sc\_V_{TI}$ .

Preliminary work (not reported) had shown a strong correlation between the relaxativity of Vistarem<sup>®</sup>,  $r_1$ , and the scaling coefficient of blood volume,  $sc\_V_B$ . To not constrain the estimation of these two parameters,  $r_{1/SVB}$  was defined as:

$$r_{1/SVB} = \frac{r_1}{sc\_V_B} \quad (21)$$

and was the parameter actually estimated.

## Statistical methods

### Bayesian framework

In a Bayesian framework (Bernardo and Smith, 1994; Gelman *et al.*, 1995), all unknowns ( $\theta$ ) are considered as random variables. Before analyzing the data, a prior probability distribution,  $P_\theta$ , is proposed for each unknown. That distribution can be interpreted in terms of degrees of beliefs about possible values of  $\theta$  and should reflect the current knowledge on  $\theta$  (from *in vitro* data, published literature, expert judgment, etc.). Informative distributions can be set for parameters whose possible values are well known (blood flows, tissue volumes, etc.). When little information is available, non-informative distributions can be used (Gelman *et al.*, 1995).

Prior distributions are then updated using new experimental data,  $y$ , to yield a posterior probability distribution. Updating is performed with Bayes' theorem, which states that posterior probability  $P(\theta|y)$  is proportional to the product of all the prior probabilities  $P_\theta$  by the data likelihood  $P(y|\theta)$ :

$$P(\theta|y) \propto \prod_{\theta} P_{\theta} \times P(y|\theta) \quad (22)$$

It is often difficult to obtain an analytical expression for  $P(\theta|y)$ , and it is the case here. Therefore, we used Markov Chain Monte Carlo (MCMC) methods to numerically obtain a sample of parameters values from their joint posterior distribution (Gelman *et al.*, 1995). Any statistical summary (*e.g.*, mean, standard deviation) can be extracted from those distributions.

To perform the analysis, the MCSim software (Bois and Maszle, 1997b) was used. For each dataset, four independent MCMC chains were run until convergence was reached (60,000 iterations). Only one on 10 in the 30,000 last iterations were recorded to produce the results presented below. For each parameter, the convergence criterion  $\hat{R}$  (Gelman *et al.*, 1995) was used to check convergence to a stable posterior

distribution. At perfect convergence, all the  $\hat{R}$  should be equal to 1. After 60,000 iterations, they were all below 1.2, so convergence was considered achieved.

#### *Prior distributions*

Unknown parameters can be classified into four groups: volumes (or corresponding scaling coefficients), plasma flow rates, diffusion rate constants and the experiment parameters (*e.g.*, time of infusion).

According to (Kohn and Melnick, 2001), blood volume represents 6% of bodyweight, so a normal distribution with mean 0.06 and standard deviation (SD) 0.006 was used for  $sc\_V_B$ .  $V_{BC}$  represents approximately 5% of the total plasma volume and the remainder is distributed between arterial and venous plasma with the ratio of 1:3 (Brown *et al.*, 1997). Therefore normal distributions with means 0.05 and 0.3 and SDs 0.005 and 0.03 were used for  $sc\_V_{BC}$  and  $sc\_V_{AP}$ , respectively. Since no prior information is available on  $sc\_V_{TC}$ , a uniform distribution between 0.1% and 20% was used.

The interstitium volume in the body corresponds approximately to 15% of bodyweight (Bischoff and Brown, 1966), so  $sc\_V_{BI}$  was *a priori* distributed lognormally with geometric mean (GM) 0.15 and geometric standard deviation (GSD) 1.2. Since interstitium volumes can be very different between tumors, the scaling parameter,  $sc\_V_{TL}$ , was distributed uniformly between 7.5% and 30%.

MRI examinations were performed on mice from day 32 to day 49 after tumor graft. Previous work (not reported) has experimentally produced a relationship between the volume of the tumor (cell line PC-3) and its age. Using it, it was possible to define a range of possible values (and a prior distribution) for the tumor volume ( $V_T$ ). As mice were imaged on different days after transplantation, four lognormal prior distributions, depending on the tumor's age, were defined. GMs were set as follow: 0.2 ml for a 32- to 34-day-old tumor, 0.4 ml for a 35-day-old tumor, 0.8 ml for a 41-day-old tumor, and 2.0 ml for 49-day-old tumor. A GSD of 1.5 was used in all cases.

Several values have been published for  $K_P$  (Brown *et al.*, 1997; Gerlowski and Jain, 1983) (with hematocrit correction). Reported values go from 0.07 ml/s to 0.14 ml/s, so a lognormal distribution with GM 0.1 ml/s and GSD 1.5 was used.

A tumor volume of 0.5 ml represents about 1.5% of the volume of a nude mouse. Therefore, as a first approximation, tumor blood flow is approximately 1.5% of the total flow (about 0.003 ml/s).  $K_{TC/T}$  was assigned a lognormal distribution with GM 0.003 ml/s and GSD of 10 (a large GSD was used as the distribution mean results from an approximation).

Since  $K_E$  is equivalent to the glomerular filtration rate (Port *et al.*, 2001b), a lognormal prior distribution with GM 0.0026 ml/s ((Gerlowski and Jain, 1983) with hematocrit correction) and GSD of 1.2 was used for that parameter.

No prior information was available for  $K_{BCI}$  or  $K_{TCI/T}$ , so uniform distributions between 0 and 0.01 were used for these parameters. It was checked that 0.01 was never reached in the simulations and is above

the plausible bound. Uniform distributions were also used for some parameters related to the experiment: the beginning and the end of infusion ( $T_{BegInf}$  and  $T_{EndInf}$ ), and  $r_{1/SVB}$ .

Table 1 sums up the prior distributions of the physiological parameters. For some parameters, truncation bounds were defined to avoid unrealistic values.

**TABLE 1:** Summary of prior and posterior distributions for the physiological parameters. Posterior distributions are given independently for treated (Group T) and untreated (Group P) mice.

Parameter	Prior Distribution	Mean	Standard deviation	Truncation bounds	Posterior distributions	
					Mean [IC 95%]	
					Group T	Group P
$sc\_V_B$	Normal	0.06	0.006	[0.02; 0.1]	0.07 [0.059; 0.081]	0.07 [0.057; 0.083]
$sc\_V_{AP}$	Normal	0.30	0.03	[0.005; 0.5]	0.14 [0.071; 0.24]	0.15 [0.075; 0.28]
$sc\_V_{BC}$	Normal	0.05	0.005	[0.01; 0.10]	0.05 [0.04; 0.06]	0.05 [0.04; 0.06]
$sc\_V_{TC}$	Uniform			[0.001; 0.2]	0.014 [0.006; 0.019]	0.021 [0.016; 0.027]
$sc\_V_{BI}$	Lognormal	0.15	1.2	[0.001; 0.2]	0.14 [0.092; 0.21]	0.14 [0.081; 0.21]
$sc\_V_{TI}$	Uniform			[0.075; 0.3]	0.20 [0.089; 0.292]	0.18 [0.077; 0.289]
$V_T^{(a)}$	Lognormal	0.2	1.5	[0.1; 0.6]	0.27 [0.13; 0.49]	0.28 [0.13; 0.50]
$V_T^{(a)}$	Lognormal	0.4	1.5	[0.2; 0.8]	0.42 [0.23; 0.68]	0.40 [0.22; 0.68]
$V_T^{(a)}$	Lognormal	0.8	1.5	[0.4; 1.6]	0.85 [0.46; 1.43]	$\_^{(b)}$
$V_T^{(a)}$	Lognormal	2.0	1.5	[1; 3]	$\_^{(b)}$	1.84 [1.10; 2.79]
$K_P$	Lognormal	0.1	1.5	[0.05; 0.3]	0.09 [0.06; 0.13]	0.11 [0.05; 0.18]
$K_{TC/T}$	Lognormal	0.003	10	[0.0003; 0.03]	0.0022 [0.0006; 0.0056]	0.0043 [0.0022; 0.0081]
$K_{BCI}$	Uniform			[0.0; 0.01]	$5.8 \times 10^{-4}$ [0.06; 22.1] $\times 10^{-4}$	$7.4 \times 10^{-4}$ [0.09; 23.6] $\times 10^{-4}$
$K_{TCI/T}$	Uniform			[0.0; 0.01]	$5.4 \times 10^{-5}$ [3.4; 9.3] $\times 10^{-5}$	$13.4 \times 10^{-5}$ [6.7; 21.6] $\times 10^{-5}$
$K_E$	Lognormal	0.0026	1.2	[0.00052; 0.013]	0.0016 [0.0010; 0.0022]	0.0017 [0.0011; 0.0026]

(a): For the tumor volume ( $V_T$ ), the prior distribution depends on the tumor age. Four distributions were then necessary.

(b): no animal were in these age groups

*Data likelihood*

Relaxation rates observed are affected by noise from different sources (recording by the MR system or automatic physiological movements). We supposed that the observed baseline corrected relaxation rates in arterial plasma,  $(\Delta R_1)_{Art}$ , and in tumor tissue,  $(\Delta R_1)_{Tum}$ , were normally distributed around actual relaxation rates ( $y_{Art}$  and  $y_{Tum}$ ) with standard deviations  $SD_{Art}$  and  $SD_{Tum}$ , respectively.

For arterial measurements (e.g., Figure 3), errors in the observed relaxation rates are approximately proportional to the values taken by those rates. Therefore,  $SD_{Art}$  was computed as:

$$SD_{Art} = CV_{Art} \times (\Delta R_1)_{Art} + (2 - CV_{Art}) \times LD_{Art} \quad (23)$$

where  $CV_{Art}$  is the coefficient of variation and  $LD_{Art}$  the detection limit (Bois *et al.*, 1994; Kashuba *et al.*, 1996). This relationship gives a  $CV_{Art}$  of 2 at the detection limit.

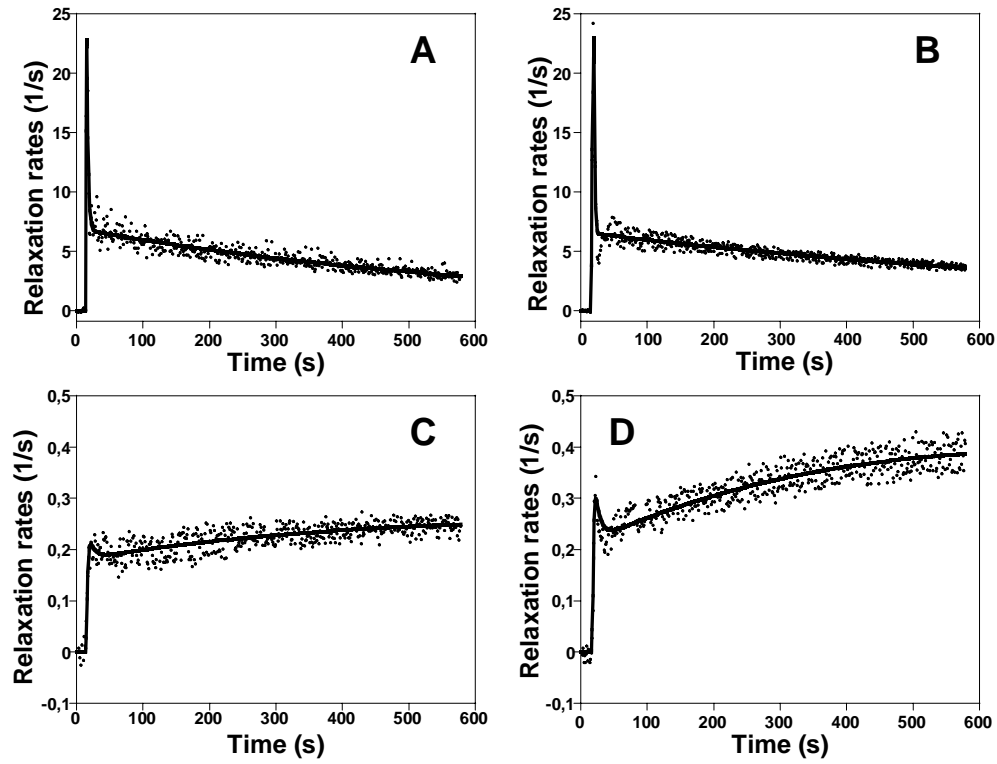
Preliminary work showed that a proportional error model is not necessary for the tumor relaxation rates. A constant error model was used.

Loguniform distributions between 0.001 and 1 were assigned to  $SD_{Tum}$ , and  $CV_{Art}$ . A non-informative uniform distribution between 0 and 0.2 was used for the parameter  $LD_{Art}$ .

**RESULTS***Model fits and posterior parameter distributions*

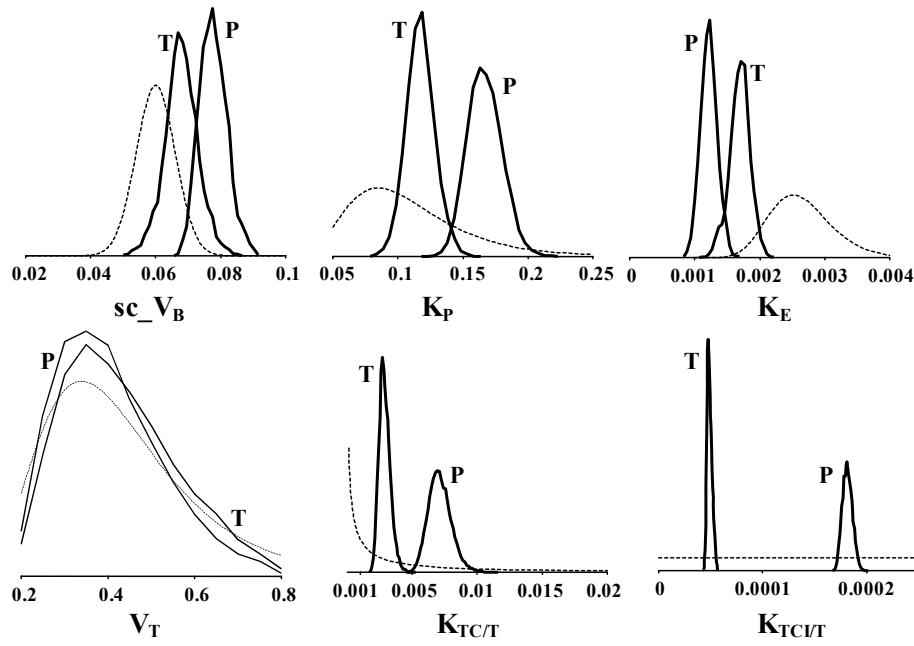
The model was fitted to the data of 13 mice. For two mice (one in each group, P1 and T6), the fitted curve and the experimental data are presented in Figure 3. The vector of parameters with the highest likelihood (the mode of the joint posterior parameter distribution) was used to produce the fits shown. The quality of fit is similar for the other mice.

Table 1 presents the posterior mean and the 95% confidence interval for treated and untreated mice. Figure 4 presents the prior and posterior distributions for some physiological parameters of a treated mouse ( $T_2$ ) and an untreated one ( $P_5$ ). These graphs illustrate two situations that can arise in parameter estimation. Information contained in the experimental data may allow an update of the prior information, and the resulting posterior distribution corresponds to a compromise, typically in favor of the data, between these two sources of information (e.g., the elimination rate constant  $K_E$ ). In such cases, parameter estimates differ according to the datasets, even if the same informative prior distribution is used (see parameter  $K_P$  in Figure 4). When little information is contained in the data for the estimation of the parameter, the posterior distribution is close to the prior (see for example, the tumor volume  $V_T$ ). Overall, physiological parameters (plasma flows, volumes ...) retained plausible values which are consistent with their prior distributions. Only the proportion of arterial plasma to total plasma (estimated to fall between 10% and 25%) was significantly lower than the published values (30% in (Brown *et al.*, 1997)).



**Figure 3:** Relaxation rates in arterial blood (A, treated mouse T6, and B, placebo mouse P1) or in tumor (C, treated mouse T6, and D, placebo mouse P1) as a function of time. The smooth curves correspond to model fits (using the parameters vector of maximum posterior density) to observed data (dots).





**Figure 4:** Graphical representations of the prior (dashed curve) and posterior (plain curves) distributions for some physiological model parameters of a treated mouse (T) and an untreated mouse (P). Parameters are: the scaling coefficient of the total blood volume ( $sc\_V_B$ ), the total plasma flow ( $K_P$ ), the elimination rate constant ( $K_E$ ), the tumor volume ( $V_T$ ), the tumor plasma flow per unit of volume ( $K_{TC/T}$ ) and the tumor permeability per unit of volume ( $K_{TC/T}$ ). The dotted curves correspond to the analytical definition of the priors (e.g., the normal distribution for  $sc\_V_B$ ), the plain curves correspond to the histogram of the Monte Carlo sampled estimates.

### ***Treatment's effects on tumor microcirculation***

Tumors' microcirculation can be quantified by: blood perfusion per unit of volume, surface permeability per unit of volume and the fraction of tumor tissue occupied by capillaries (tissue fractional blood volume). Blood perfusion per unit of tumor volume is given by the parameter  $K_{TC/T}$  divided by (1 - hematocrit), the surface permeability coefficient corresponds directly to the parameter  $K_{TC/T}$  and the fraction of capillaries in tumor is obtained by dividing parameter  $sc\_V_{TC}$  by (1 - hematocrit). For each mouse, the average estimate and the 95% confidence interval (CI) of those microcirculation' parameters are represented on Figure 5.

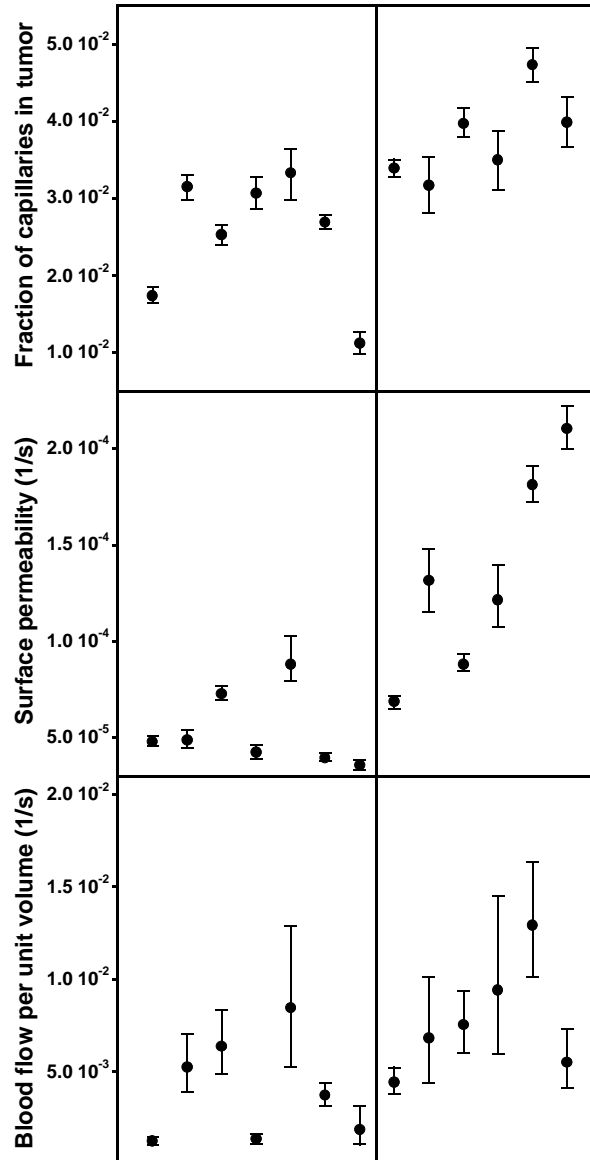
For each group of mice, individual mouse parameter vectors were pooled to obtain group averages (so 72,000 parameter vectors were used). For treated mice, the mean of tumor blood flow per unit of volume is  $0.0040 \text{ s}^{-1}$  and its CI is  $[0.0011 - 0.01024]$ . For placebo mice (Group P), the mean is  $0.0078 \text{ s}^{-1}$  with CI  $[0.0040 - 0.0147]$ . The means of the surface permeability per unit of volume were respectively  $5.38 \times 10^{-5} \text{ s}^{-1}$  (CI:  $[3.45 \times 10^{-5}; 9.3 \times 10^{-5}]$ ) and  $13.38 \times 10^{-5} \text{ s}^{-1}$  (CI:  $[6.71 \times 10^{-5}; 21.64 \times 10^{-5}]$ ) for treated and placebo mice. The

proportions of capillary volume in tumor were respectively estimated to 2.5% (CI: [1.0; 3.5]) and to 3.8% (CI: [3.0; 4.8]) for treated and placebo mice.

To test the difference between these two groups, we used a Monte Carlo method. We computed differences between placebo and treated group parameters after randomly pairing parameters vectors from each group. To obtain a reliable estimate of the difference, pairing was randomized one thousand times. On average, the probability for a treated mouse to have a tumor blood flow higher than for an untreated mouse is 0.183. For the permeability per unit of volume and the fraction of capillary volume, these probabilities are respectively 0.057 and 0.053. A one-way ANOVA for these parameters (estimates used were the modes of the joint posterior distribution for each mouse) confirmed these results.

To test the global effect of ZD4190 treatment on tumor's microcirculation, a multivariate ANOVA of the three microcirculation parameters for the factor treatment was performed. The estimates used were also the modes of the joint posterior distribution. Both Pillai's trace and Hotelling's trace tests indicated a significant difference between the two groups of mice ( $p = 0.029$ ).

No significant difference was found between ZD4190 treated and placebo mice for the whole-body physiological parameter (*e.g.*, total plasma flow, plasma volume *etc.*). We may conclude from these results that ZD4190 treatment has no significant effect on whole-body circulation, but has a beneficial action on reducing tumor's microcirculation.



**Figure 5:** *Posterior distribution of the microcirculation parameters for each of the 13 mice. Posterior means are represented by black dots, the limits are respectively the 2.5th and the 97.5th percentile. Treated mice are to the left of the separation line and placebo mice to the right.*

### **Comparison with a 2-compartment model**

To compare the results obtained with our whole-body model and those obtained with a simple pharmacokinetic model, we analyzed the datasets of four mice with the 2-compartment model proposed by Pradel *et al.* (Pradel *et al.*, 2003). The latter describes tumor kinetics and an arterial input. These two models differ in some points. In (Pradel *et al.*, 2003), the arterial data are directly used as input in a tumor PK model (which implies no uncertainty in these data), the measurement errors are fixed (or known with precision), and exchanges between compartments are mono-directional. Table 2 presents the microcirculation parameter estimates obtained with the two models.

**TABLE 2:** Comparison of the tumor microcirculation estimates obtained with our PBPK model and a two-compartment model (Pradel *et al.*, 2003). *Perfusion*, *Permeability* and *Capillaries* stand respectively for blood perfusion per unit of volume, surface permeability per unit of volume and the fraction of tumor tissue occupied by capillaries.

Mouse	Whole body PBPK model			2-compartment model		
	<i>Perfusion</i> ( $10^{-3} \text{ s}^{-1}$ )	<i>Permeability</i> ( $10^{-5} \text{ s}^{-1}$ )	<i>Capillaries</i> (%)	<i>Perfusion</i> ( $10^{-3} \text{ s}^{-1}$ )	<i>Permeability</i> ( $10^{-5} \text{ s}^{-1}$ )	<i>Capillaries</i> (%)
T <sub>1</sub>	1.2 [1.0; 1.5]	4.8 [4.6; 5.1]	1.7 [1.6; 1.8]	1.3 [1.1; 1.6]	4.5 [4.4; 4.6]	1.8 [1.7; 1.9]
T <sub>6</sub>	3.7 [3.2; 4.4]	4.0 [3.8; 4.2]	2.7 [2.6; 2.8]	3.6 [3.3; 4.2]	3.7 [3.6; 3.8]	2.7 [2.6; 2.8]
P <sub>1</sub>	4.4 [3.8; 5.2]	6.9 [6.5; 7.2]	3.4 [3.3; 3.5]	4.3 [3.8; 5.3]	5.4 [5.3; 5.5]	3.7 [3.6; 3.8]
P <sub>3</sub>	7.5 [6.0; 9.4]	8.8 [8.5; 9.4]	4.0 [3.8; 4.2]	8.1 [7.0; 9.2]	8.0 [7.8; 8.2]	4.1 [3.9; 4.2]

## DISCUSSION

The purpose of this work was first to model the distribution of a rapid clearance blood pool agent Vistarem<sup>®</sup>, in the body and a grafted tumor, and to characterize the tumor microcirculation. The model was then applied to evaluate changes in tumor microcirculation after treatment of mice with an antiangiogenic drug, ZD4190. Our model was analyzed in a Bayesian framework in order to avoid identifiability problems and to measure the uncertainty on the estimated parameters. The physiological model is quite general and can be applied to agents exhibiting a physiological behavior similar to that of Vistarem<sup>®</sup> (limited interstitial diffusion, free renal excretion).

For Vistarem<sup>®</sup>, our results show that, in nude mice, the average estimate of its elimination rate constant,  $K_E$ , is 0.0017 ml/s (with individual estimates falling between 0.0010 ml/s and 0.0024 ml/s for treated or placebo animals). For a RCBPA,  $K_E$  is supposed to be "equivalent" to the glomerular filtration rate (GFR). Our estimate is approximately 35% lower than the published value for GFR (0.0026 ml/s) (Gerlowski and Jain, 1983; Singer, 2001), which was also our prior mean. Either nude mice have lower GFR than standard mice, or Vistarem<sup>®</sup> is eliminated at a slightly lower rate than the GFR. Nonetheless, Vistarem<sup>®</sup> can still be considered as a rapid clearance agent.

Tumor microcirculation was then quantified in a tumor region using our approach. The tumor perfusion, permeability and fraction of capillaries were only evaluated in that region. To take that into account the fact that tumors are heterogeneous, the tumor could be split into several regions, in which measurements could be performed. Adjustment of the pharmacokinetic model to individual regions could provide parametric maps of the parameters of interest (Ludemann *et al.*, 2000). Care should be taken to model interactions between neighboring tumor regions.

In this work, measurements were performed in arterial blood and tumor to better characterize the average tumor microcirculation parameters (Port *et al.*, 2001c). To describe such measurements, compartmental models describing tumor kinetics and an arterial input have been proposed. Either the arterial data are directly used as input in a tumor PK model (Pradel *et al.*, 2003), or they are modeled independently by a tri-exponential function (Daldrup *et al.*, 1998). We adjusted Pradel's model (Pradel *et al.*, 2003) to our data for four few mice. The results obtained for the tumor microcirculation parameters are of the order of those obtained with our whole-body PBPK model. The slight differences can be explained by the fact that Pradel's model neglects transfers from tumor interstitium to tumor capillaries, and that it does not account for uncertainty in the arterial measurements. Here, we developed a whole-body model based on physiology in which tissues are interconnected through the body fluid system. Twelve parameters were necessary to describe the Vistarem<sup>®</sup> distribution, three model the uncertainty in input (injection), and the last three the measurement errors. Only three parameters are needed in (Pradel *et al.*, 2003), and nine in (Daldrup *et al.*, 1998). In (Daldrup *et al.*, 1998), injection is not modeled and the measurement errors are fixed. Because our parameters, describing the time course of the Vistarem<sup>®</sup> in arterial blood, have a physiological meaning, informative prior distribution can be used for them, which further reduces the difference in terms of complexity/information ratio.

Our results show that compartmental and physiological models provide similar fits for the same datasets. Yet, those two kinds of model differ in their aims. The choice of structure for compartmental models depends only on their fit to the data. Physiological models are defined on the basis of the body anatomy and make use of *a priori* information from the literature, previous or ancillary experiments to describe the fate of the substance administered. In addition, physiological models, first fitted with vascularization data, can then be used to describe the fate of different drugs. That is beyond the scope of tumor vascularization studies, but has general implications for the development and use of such models in the therapeutic drug industry.

Physiological modeling provides the opportunity to use relevant prior information on parameter values. To take into account the fact that prior information and data are affected by uncertainty, our analysis was conducted in the Bayesian framework. Defining prior distributions for physiological parameters required literature searches and the making of some assumptions. Even though it is well-known that these parameters exhibit a wide range of inter-individual variability, the values available are for a "reference mouse" usually without mention of ranges or other measures of variability or uncertainty. Values for a particular strain of mice (here, Nude) or anaesthetized mice are also difficult to find. To overcome these difficulties, we used reference values to define the mean of the prior distributions, and we set standard deviations and truncation limits to reasonable values allowing for variability among mice. As shown in Figure 4, variability in posterior distributions was allowed, even though the same prior distributions were used.

A problem for physiological model specification is the degree of complexity to adopt. Four compartments (and two subcompartments) were used here to describe Vistarem<sup>®</sup> kinetics. The model fit is quite good for arterial plasma and tumor data. Yet the model was not able to reproduce the small

recirculation peak apparent in the arterial blood data. A remedy could be, for example, the inclusion of a diffusion model (Lafrance *et al.*, 2002; Oliver *et al.*, 2001). However parameter identifiability problems could arise with a more complex model.

In typical pharmacokinetics studies, concentrations in the body fluids (blood or urine, for example) at a few points in time are often the only data available. One of the major promises of imaging techniques, such as MRI, is to provide numerous measurements of agent's concentrations in any part of the body without sacrificing the subject. Moreover, the resulting MR images can inform on the anatomy of the subject, and give, for example, indications on the organ sizes or blood flows, which are typical physiological model parameters. That is why it is particularly interesting to couple the physiological modeling with imaging technology. Modeling can offer a global description of imaging data, amenable to statistical validation and providing for improved extrapolations in times, doses, administration route, or transposition across species (Bernillon and Bois, 2000; Watanabe and Bois, 1996). In turn, the availability of internal concentration data at short-term intervals can offer unique opportunities to refine and improve pharmacokinetic models.

## REFERENCES

- Bernardo, J. M., and Smith, A. F. M. (1994). Bayesian Theory. Wiley, New York.
- Bernillon, P., and Bois, F. Y. (2000). Statistical issues in toxicokinetic modeling: A Bayesian perspective. *Environmental Health Perspectives* 108, 883-893.
- Bischoff, K. B., and Brown, R. G. (1966). drug distribution in mammals. *Chemical Engineering Progress* 66, 33-45.
- Bois, F. Y., and Maszle, D. (1997). MCSim: a simulation program. *Journal of Statistical Software* 2(9), <http://www.stat.ucla.edu/journals/jss/v02/i09> (software available online at [www.toxi.ineris.fr](http://www.toxi.ineris.fr)).
- Bois, F. Y., Tozer, T. N., Hauck, W. W., Chen, M. L., Patnaik, R., and Williams, R. L. (1994). Bioequivalence: performance of several measures of extent of absorption. *Pharmaceutical Research* 11, 715-22.
- Brix, G., Semmler, W., Port, R., Schad, L. R., Layer, G., and Lorenz, W. J. (1991). Pharmacokinetic parameters in CNS Gd-DTPA enhanced MR imaging. *Journal of Computer Assisted Tomography* 15, 621-8.
- Brown, R. P., Delp, M. D., Lindstedt, S. L., Rhomberg, L. R., and Beliles, R. P. (1997). Physiological parameter values for physiologically based pharmacokinetic models. *Toxicology and Industrial Health* 13, 407-484.
- Daldrup, H. E., Shames, D. M., Hussein, W., Wendland, M. F., Okuhata, Y., and Brasch, R. C. (1998). Quantification of the extraction fraction for gadopentetate across breast cancer capillaries. *Magnetic Resonance in Medicine* 40, 537-543.
- Davis, B., and Morris, T. (1993). Physiological parameters in laboratory animals and human. *Pharmaceutical Research* 10, 1093-1095.

Gelman, A., Bois, F. Y., and Jiang, J. (1996). Physiological pharmacokinetic analysis using population modeling and informative prior distributions. *Journal of the American Statistical Association* 91, 1400-1412.

Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (1995). *Bayesian Data Analysis*. Chapman & Hall, London.

Gerlowski, L. E., and Jain, R. K. (1983). Physiologically based pharmacokinetic modeling: principles and applications. *Journal of Pharmaceutical Sciences* 72, 1103-1127.

Gillies, R. J., Bhujwala, Z. M., Evelhoch, J., Garwood, M., Neeman, M., Robinson, S. P., Sotak, C. H., and Van Der Sanden, B. (2000). Applications of magnetic resonance in model systems: tumor biology and physiology. *Neoplasia* 2, 139-151.

Kashuba, A. D., Ballow, C. H., and Forrest, A. (1996). Development and evaluation of a Bayesian pharmacokinetic estimator and optimal, sparse sampling strategies for ceftazidime. *Antimicrobial Agents and Chemotherapy* 40, 1860-5.

Kohn, M. C., and Melnick, R. L. (2001). Physiological modeling of butadiene disposition in mice and rats. *Chemico-Biological Interactions* 135-136, 285-301.

Lafrance, P., Lemaire, V., Varin, F., Donati, F., Belair, J., and Nekka, F. (2002). Spatial effects in modeling pharmacokinetics of rapid action drugs. *Bulletin of Mathematical Biology* 64, 483-500.

Ludden, T. M., Gillespie, W. R., and Bachman, W. J. (1995). Commentary on "Physiologically based pharmacokinetic modeling as a tool for drug development". *Journal of Pharmacokinetics and Biopharmaceutics* 23, 231-235.

Ludemann, L., Hamm, B., and Zimmer, C. (2000). Pharmacokinetic analysis of glioma compartments with dynamic Gd-DTPA-enhanced magnetic resonance imaging. *Magn Reson Imaging* 18, 1201-14.

Oliver, R. E., Jones, A. F., and Rowland, M. (2001). A whole-body physiologically based pharmacokinetic model incorporating dispersion concepts: Short and long time characteristics. *Journal of Pharmacokinetics and Pharmacodynamics* 28, 27-55.

Port, M., Corot, C., Raynal, I., Idee, J. M., Dencausse, A., Lancelot, E., Meyer, D., Bonnemain, B., and Lautrou, J. (2001a). Physicochemical and biological evaluation of P792, a rapid- clearance blood-pool agent for magnetic resonance imaging. *Investigative Radiology* 36, 445-454.

Port, M., Corot, C., Rousseaux, O., Raynal, I., Devoldere, L., Idee, J.-M., Dencausse, A., Greneur, S. L., Simonot, C., and Meyer, D. (2001b). P792: a rapid clearance blood pool agent for magnetic resonance imaging: preliminary results. *Magnetic Resonance Materials in Physics Biology and Medicine* 12, 121-127.

Port, R. E., Knopp, M. V., and Brix, G. (2001c). Dynamic contrast-enhanced MRI using Gd-DTPA: interindividual variability of the arterial input function and consequences for the assessment of kinetics in tumors. *Magn Reson Med* 45, 1030-8.

Port, R. E., Knopp, M. V., Hoffmann, U., Milker-Zabel, S., and Brix, G. (1999). Multicompartment analysis of gadolinium chelate kinetics: blood-tissue exchange in mammary tumors as monitored by dynamic MR imaging. *Journal of Magnetic Resonance Imaging* 10, 233-241.

Pradel, C., Siauve, N., Bruneteau, G., Clement, O., de Bazelaire, C., Frouin, F., Wedge, S. R., Tessier, J. L., Robert, P. H., Frija, G., and Cuenod, C. A. (2003). Reduced capillary perfusion and permeability in human tumour xenografts treated with the VEGF signalling inhibitor ZD4190: an in vivo assessment using dynamic MR imaging and macromolecular contrast media. *Magnetic Resonance Imaging* 21, 845-51.

Singer, M. A. (2001). Of mice and men and elephants: metabolic rate sets glomerular filtration rate. *American Journal of Kidney Diseases* 37, 164-178.

Taylor, J. S., Tofts, P. S., Port, R., Evelhoch, J. L., Knopp, M., Reddick, W. E., Runge, V. M., and Mayr, N. (1999). MR imaging of tumor microcirculation: promise for the new millennium. *Journal of Magnetic Resonance Imaging* 10, 903-907.

Tofts, P. S. (1997). Modeling tracer kinetics in dynamic Gd-DTPA MR imaging. *Journal of Magnetic Resonance Imaging* 7, 91-101.

Tofts, P. S., Berkowitz, B., and Schnall, M. D. (1995). Quantitative analysis of dynamic Gd-DTPA enhancement in breast tumors using a permeability model. *Magnetic Resonance in Medicine* 33, 564-8.

Tofts, P. S., and Kermode, A. G. (1991). Measurement of the blood-brain barrier permeability and leakage space using dynamic MR imaging. 1. Fundamental concepts. *Magnetic Resonance in Medicine* 17, 357-67.

Watanabe, K., and Bois, F. Y. (1996). Interspecies extrapolation of physiological pharmacokinetic parameter distributions. *Risk Analysis* 16, 741-754.

Wedge, S. R., Ogilvie, D. J., Dukes, M., Kendrew, J., Curwen, J. O., Hennequin, L. F., Thomas, A. P., Stokes, E. S. E., Curry, B., Richmond, G. H. P., and Wadsworth, P. F. (2000). ZD4190: An orally active inhibitor of vascular endothelial growth factor signaling with broad-spectrum antitumor efficacy. *Cancer Research* 60, 970-975.



## **6. Utilisation de traceurs non-toxiques dans les études toxicocinétiques**

Dans cette section, nous nous intéressons à l'utilisation des traceurs dans les études toxicocinétiques. Notre travail se divise en deux. Dans une première partie, nous étudions le bénéfice apporté par une exposition à un traceur non-toxique, en complément d'une exposition à une substance toxique, et proposons des protocoles d'expositions à ces deux substances permettant de réduire la dose toxique reçue par le volontaire. Une des limites de ces protocoles est que les systèmes mathématiques résultants peuvent être de grande dimension. Nous traitons ce problème dans la seconde partie de cette section, et proposons d'appliquer des techniques d'agrégation (ou réduction) de système d'équations différentielles.

## **Use of a chemical probe to increase safety for human volunteers in toxicokinetic studies**

Céline Brochot<sup>1</sup> and Frédéric Y. Bois<sup>1</sup>

<sup>1</sup>INERIS, Institut National de l'Environnement Industriel et des Risques, Unité de Toxicologie Expérimentale, Parc Alata BP2, 60550 Verneuil En Halatte, France ([celine.brochot@ineris.fr](mailto:celine.brochot@ineris.fr), [frederic.bois@ineris.fr](mailto:frederic.bois@ineris.fr))

Soumis à *Risk Analysis*

## RESUME

Afin d'éviter les extrapolations inter-espèces dans les études toxicocinétiques et lors du développement des substances thérapeutiques, il est recommandé de développer directement des données chez l'humain. Dans ce cas, la dose à laquelle les volontaires sont exposés, ne doit engendrer qu'un risque négligeable, mais doit être suffisamment élevé pour permettre la quantification de la substance dans les tissus. Nous proposons de réduire la dose reçue par les volontaires humains lors de l'exposition, et de compenser la perte d'information en exposant les mêmes volontaires à un agent non-toxique.

Cette méthode a été appliquée pour développer des protocoles d'exposition pour le 1,3-butadiène chez l'humain. Pour valider notre procédure, nous avons travaillé avec des données simulées. Trois temps d'exposition (20, 10 et 5 minutes) et quatre concentrations d'exposition (2, 1, 0.8, et 0.5 ppm) ont été utilisés pour définir 12 scénarios d'exposition pour le butadiène. L'isoflurane a été utilisé comme traceur, avec une exposition simulée de 20 ppm d'isoflurane pendant 15 minutes. Les concentrations d'isoflurane et de butadiène dans l'air exhalé ont été supposées être mesurées 5 fois durant l'exposition et 5 fois après. Un modèle physiologique toxicocinétique à trois compartiments a été utilisé pour décrire conjointement les cinétiques de l'isoflurane et du butadiène. Pour chaque sujet, les données concernant le butadiène ont été analysées dans un contexte bayésien, avec ou sans les données d'isoflurane. Lorsque les données butadiène et isoflurane sont considérées conjointement, la précision de la constante de métabolisme du butadiène ou de la fraction métabolisée est augmentée, et le biais réduit.

Une exposition à 10 ppm×min de butadiène et à 300 ppm×min d'isoflurane donne une précision et un biais équivalents à une unique exposition à 40 ppm×min de butadiène. En conséquence, la dose de butadiène reçue par les volontaires pourrait être divisée par 4, si l'exposition au butadiène était complétée d'une exposition à un traceur non-toxique. Cette approche est très générale, et pourrait être utilisée pour d'autres substances chimiques.

**Mots Clés :** modélisation PBPK; traceur chimique; 1,3-butadiène; expositions humaines.

**ABSTRACT**

To avoid inter-species extrapolation in toxicokinetics and drug development, it is convenient to directly develop human data. In that case, exposure dose should pose null or negligible risk to the exposed individual, but still be sufficiently high to allow quantification. We propose to reduce the dose received by human volunteers during exposure, and to compensate for loss of information by exposing the same volunteers to a non-toxic agent. This method was applied to develop 1,3-butadiene (BD) exposure protocols for humans. To validate our procedure, we worked with simulated data. Three exposure times and four exposure concentrations were used to define 12 inhalation exposure scenarios for BD. Isoflurane was used as probe. Isoflurane or BD exhaled air concentrations were supposed to be measured 5 times during exposure and 5 times afterward. A three-compartment physiological toxicokinetic model was used to jointly describe BD and isoflurane data. For each subject, BD data were analyzed, in a Bayesian framework, either alone or together with the isoflurane data. The precision of BD metabolic rate constant or fraction metabolized was increased, and bias reduced, when BD and probe data were considered jointly. An exposure to 10 ppm×min BD and 300 ppm×min isoflurane gave equivalent precision and bias as a unique exposure to 40 ppm×min BD. The BD dose received by volunteers could therefore be at least quartered, if BD exposure was supplemented with that to a probe. This approach is fairly general and could be used for many other chemicals.

**Keywords:** PBPK modeling; chemical probe; 1,3-butadiene; human exposures.

## INTRODUCTION

The health risk incurred during an exposure to a chemical is not simply related to the exposure dose. Indeed, the effect intensity depends essentially on the “effective dose” (quantity or concentration) of the active substance reaching tissues or cells (Andersen, 1995a). Such amounts result from complex processes involving absorption (A), distribution (D), metabolism (M) and excretion (E) of the substance. These ADME processes are affected by the anatomy and physiology of the exposed subject (organ sizes, blood perfusion rates...), but also age, sex or physical activity levels etc. (Rowland and Tozer, 1989). Toxicokinetic modeling aims at understanding and describing the link between exposure to a substance and its effective dose(s). Physiologically based pharmacokinetic (PBPK) models are popular tools in this field. These models are defined on the basis of the body anatomy and use physiological parameters (volumes, flows...) to describe the fate of the substance administered.

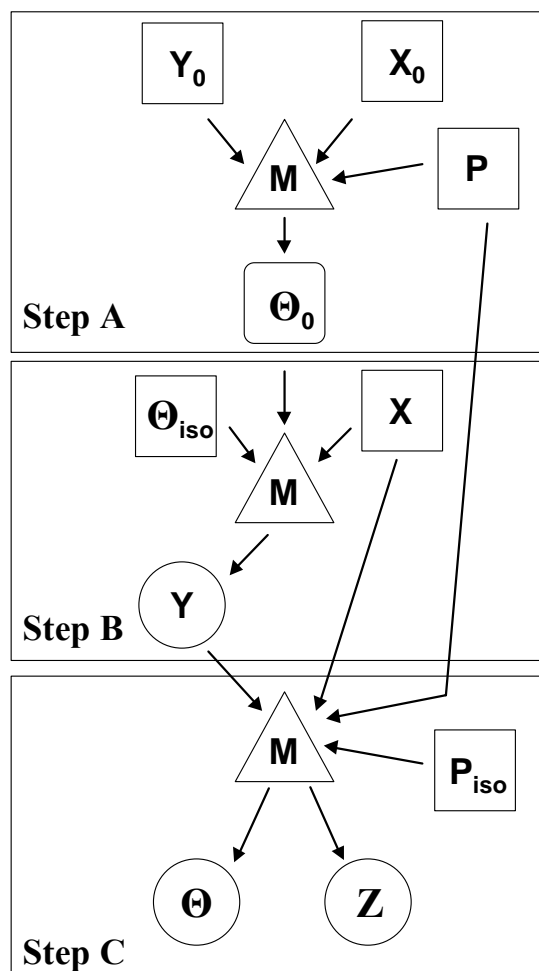
Typically, the experimental data available for the development and validation of toxicokinetic models are concentrations of the substance in blood, urine, or tissues on experimentally exposed animals. In this case, inferring ADME processes in humans requires inter-species transpositions. In case of PBPK modeling, those can be performed by scaling the model's parameters (such as volumes or blood flows) to the species of interest (Bernillon and Bois, 2000; Watanabe and Bois, 1996). Model predictions can be obtained with that new set of "human" parameters. However such transpositions cause problems if the chemical reaches organs that have no correspondence between species (Bailer and Dankovic, 1997) or if wide variations in a process are suspected between species – *e.g.*, for 1,3-butadiene metabolism (Himmelstein *et al.*, 1996), or for pesticides toxicity (NRC, 2004). In that case, without species-specific data, it is impossible to adequately characterize and quantify the toxicokinetics of a substance.

Consequently, for such difficult problems, experimental studies are conducted in which human volunteers receive controlled low-level exposures of a toxic substance. Those exposures should, obviously, pose null or negligible risk to the volunteers. A main concern is therefore to reduce, as much as possible, the dose and risk received by the subjects. On the other hand, exposure should still be sufficiently high to allow quantification in biological fluid or tissues. In this work, we propose to lower the exposure dose of the toxic substance and compensate for loss of information by exposing the same subjects to a non-toxic agent. If the pharmacokinetics of the toxic substance and the probe can be modeled by the same anatomical parameters, data collected from probe exposure can give information on the anatomy and physiology of study exposed subjects, whereas data collected from toxic exposure inform on the toxicokinetic parameters specific to the substance of interest.

As an illustration, we use this methodology to develop 1,3-butadiene exposure protocols for humans, in order to quantify butadiene metabolism. Butadiene is a gas (under normal conditions) used in the production of synthetic rubber and plastics, and it has also been detected in cigarette smoke and automobile emission. It is carcinogenic for some laboratory animals, and is suspected to be a human carcinogen

(International Agency for Research on Cancer, 1992). For this exploratory study, the non-toxic substance chosen was isoflurane (an inhalation anaesthetic gas).

Different experimental designs for butadiene and isoflurane exposures were tested. In order to validate our protocol, we used simulated data. Figure 1 presents the logic followed during this study. The work is subdivided in three steps: inferring the physiological characteristics of individuals in a human population from actual butadiene exposures (Step A), simulating "experimental data" for a set of subjects (Step B), and adjusting a model to these fake data and obtaining model predictions (Step C).



**Figure 1:** A logical diagram of our study. The work was subdivided in three steps: inferring the physiological characteristics of individuals in a human population from actual butadiene exposures (Step A), simulating "experimental data" for a set of subjects (Step B), and adjusting a model to these fake data and obtaining model predictions (Step C). Step A was performed in a previous study (Mezzetti et al., 2003). Symbols  $X$ ,  $Y$ ,  $\Theta$  and  $Z$  represent respectively the experimental design (exposure concentrations and times, and the sampling times), the data (including the covariates), the parameter estimates and the model's predictions.  $P$  and  $P_{iso}$  are respectively the prior distributions for the butadiene and physiological parameters, and for the isoflurane parameters.  $\Theta_{iso}$  are the values of isoflurane parameters used for the data simulation. The symbol  $M$  in a triangle corresponds to the three-compartment PBPK model.

## METHODS

In this section, we present the three-compartment PBPK model we used for 1,3-butadiene and isoflurane, the exposure scenarios we tested, the generation of the data, the Bayesian calibration of the model

and methods used to evaluate exposure scenarios. In order to be realistic, the study conducted by Mezzetti *et al.* (2003) was selected as a basis to define our experimental conditions. They studied 1,3-butadiene metabolism in humans, with inhalation exposure to 2 parts per million (ppm) of butadiene for 20 minutes.

### **Physiological toxicokinetic model (Symbol *M* in Figure 1)**

A three-compartment PBPK model, depicted Figure 2, was used to describe butadiene (Mezzetti *et al.*, 2003; Smith *et al.*, 2001) and isoflurane kinetics. The compartments represent the well-perfused (*wp*) tissues, the poorly perfused (*pp*) tissues and fat. They are connected in parallel by the blood circulatory system. This model assumes that metabolism takes place in the well-perfused tissues and describes it by a first order rate ( $k_{met}$ ), adequate for low butadiene exposure simulations (Bois *et al.*, 1999b). The time course of the quantity of butadiene in each compartment is given by the following set of first order differential equations:

$$\begin{aligned}\frac{\partial Q_{wp}(t)}{\partial t} &= F_{wp} \times \left( C_{art}(t) - \frac{Q_{wp}(t)}{V_{wp} \times PC_{wp}} \right) - k_{met} \times Q_{wp}(t) \\ \frac{\partial Q_{pp}(t)}{\partial t} &= F_{pp} \times \left( C_{art}(t) - \frac{Q_{pp}(t)}{V_{pp} \times PC_{pp}} \right) \\ \frac{\partial Q_{fat}(t)}{\partial t} &= F_{fat} \times \left( C_{art}(t) - \frac{Q_{fat}(t)}{V_{fat} \times PC_{fat}} \right)\end{aligned}\quad (1)$$

The following notations are used for each compartment *i*:  $Q_i$ , the substance amount;  $V_i$ , the volume;  $F_i$ , the blood flow; and  $PC_i$ , the tissue over blood partition coefficient. The term  $C_{art}$  stands for the arterial blood concentration. The cumulated amount metabolized,  $Q_{met}$ , is given by:

$$\frac{\partial Q_{met}(t)}{\partial t} = k_{met} \times Q_{wp}(t) \quad (2)$$

To deal with inhalation exposures to butadiene and isoflurane, lungs were split into two sub-regions: alveolar space (where gas exchanges occur with the ambient air) and the airways (also called dead-space). The ratio of dead-space volume over total lungs volume is the dilution factor,  $f_{DS}$ . Assuming instantaneous equilibrium of the concentrations at the lung level, the concentration in the alveolar region can be easily related to the arterial blood concentration through the blood-air partition coefficient  $PC_{a/b}$ . The concentration in exhaled breath is then expressed as a mixture of arterial blood and inhaled concentrations:

$$C_{exh}(t) = (1 - f_{DS}) \times \frac{C_{art}(t)}{PC_{a/b}} + f_{DS} \times C_{inh}(t) \quad (3)$$

The total blood flow ( $F_{tot}$ ) is related to the alveolar ventilation rate ( $F_{alv}$ ) through the ventilation perfusion rate ( $VPR$ ). The alveolar ventilation rate is also linked to the pulmonary ventilation rate by the dilution factor. Let denote the quantity of butadiene coming out of each of the three compartments by  $dQ_{ven}$ . We have the following relationships:



$$\begin{aligned}
F_{tot} &= \frac{F_{alv}}{VPR} \\
F_{alv} &= F_{pul} \times (1 - f_{DS}) \\
\frac{\partial Q_{ven}(t)}{\partial t} &= F_{fat} \frac{Q_{fat}(t)}{V_{fat} \times PC_{fat}} + F_{wp} \frac{Q_{wp}(t)}{V_{wp} \times PC_{wp}} + F_{pp} \frac{Q_{pp}(t)}{V_{pp} \times PC_{pp}}
\end{aligned} \tag{4}$$

At equilibrium, the arterial blood concentration is then given by:

$$C_{art}(t) = \frac{F_{alv} \times C_{inh}(t) + \partial Q_{ven}(t) / \partial t}{\frac{F_{alv}}{PC_{a/b}} + F_{tot}} \tag{5}$$

The same model is used to describe isoflurane kinetics. The only difference concerns partition coefficient ( $PC_i$ ) and metabolic rate ( $k_{met}$ ) values. Scientific literature suggests that quantity of isoflurane metabolized after inhalation exposure is around 0.2% (Carpenter *et al.*, 1986; Haeberle *et al.*, 2002). Isoflurane metabolism was therefore neglected and its  $k_{met}$  was set to zero. To describe butadiene kinetics together with isoflurane kinetics, two systems of three differential equations each are then necessary. These two systems have common physiological parameters which are subject-specific and only differ in the values for partition coefficients and metabolic rate.

Compartments' effective volumes were expressed as a function of the subject's body mass ( $BM$ ) and fraction of body mass corresponding to the given tissue ( $sc\_V_i$ ):

$$V_i = sc\_V_i \times BM \tag{6}$$

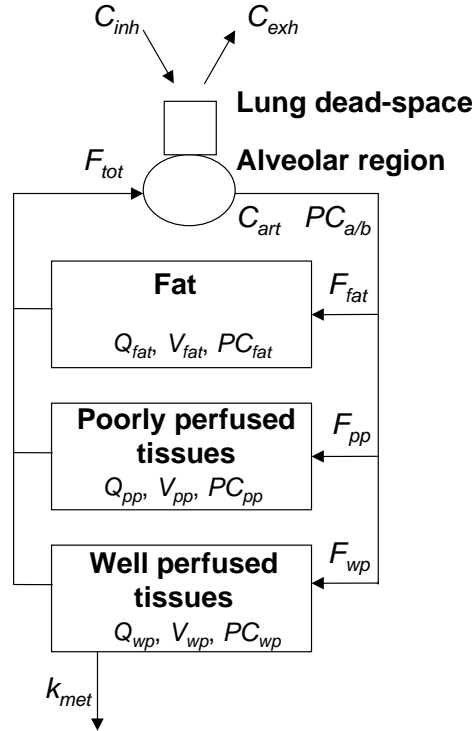
The relative fat volume of each subject was computed from a formula developed by Deurenberg *et al.* (1991) based on height, body mass, age and sex:

$$sc\_V_{fat} = \frac{1.2 \times \frac{BM}{Height^2} - 10.8 \times (2 - Sex) + 0.23 \times Age - 5.4}{100} \tag{7}$$

where  $Sex$  is 1 for males and 2 for females. Approximating density to 1 for all tissues, we constrained the sum of all the compartment volumes to 90% of body mass (the remaining 10% corresponds to bones). The poorly perfused tissues volume was expressed as the residual mass remaining after the well perfused and fat fractions were subtracted from 0.9. A similar approach was used for blood flow in each compartment. Blood flows of fat and poorly perfused tissues were scaled to the total blood flow ( $F_{tot}$ ):

$$F_i = sc\_F_i \times F_{tot} \tag{8}$$

Since the sum of blood flows is constrained by the total blood flow, blood flow in well perfused tissues was deduced by subtracting  $F_{pp}$  and  $F_{fat}$  from  $F_{tot}$ .



**Figure 2:** The three-compartment PBPK model used to describe butadiene (or isoflurane) kinetics in humans. Each compartment  $i$  is characterized by the substance quantity ( $Q_i$ ), volume ( $V_i$ ), tissue over blood partition coefficient ( $PC_i$ ), and blood flow ( $F_i$ ). Metabolism (modeled as first order, with rate constant  $k_{met}$ ) only occurs in well perfused tissues (viscera). Pulmonary exchanges are governed by the pulmonary ventilation rate and the blood-air partition coefficient ( $PC_{a/b}$  for butadiene and  $PC'_{a/b}$  for isoflurane).

For butadiene exposure, the set of parameters is therefore given by:

$$\theta = \{sc\_V_{wp}, sc\_V_{fat}, sc\_F_{pp}, sc\_F_{fat}, F_{pul}, f_{DS}, VPR, PC_{a/b}, PC_{pp}, PC_{wp}, PC_{fat}, k_{met}\} \quad (9)$$

For subjects exposed to both butadiene and isoflurane, this set is increased with the four isoflurane partition coefficients ( $PC'_i$ ):

$$\theta' = \{sc\_V_{wp}, sc\_V_{fat}, sc\_F_{pp}, sc\_F_{fat}, F_{pul}, f_{DS}, VPR, PC_{a/b}, PC_{pp}, PC_{wp}, PC_{fat}, k_{met}, PC'_{a/b}, PC'_{pp}, PC'_{wp}, PC'_{fat}\} \quad (10)$$

### The simulated human population (Step A in Figure 1)

Mezzetti *et al.* (2003) have studied 1,3-butadiene metabolism in 133 humans. Subjects were exposed to 2 ppm of butadiene for 20 minutes (design  $X_0$  in Figure 1). For each subject, concentration of butadiene in exhaled air was measured at ten times, 5 during exposure and 5 afterward (data  $Y_0$  in Figure 1). In addition to these concentrations, each subject was defined by its age, sex, height, weight and pulmonary ventilation rate ( $F_{pul}$ ), butadiene blood air partition coefficient ( $PC_{a/b}$ ) and the dilution factor ( $f_{DS}$ ). Mezzetti *et al.* (2003)

adjusted the three-compartment PBPK model to the 133 human datasets in a Bayesian framework (with prior distributions  $P$  in Figure 1), a posterior distribution was then estimated for each model's parameter for each subject.

For our purposes, values for the parameter vector  $\theta'$  need to be set to simulate butadiene and isoflurane concentrations in exhaled air using the above three-compartment PBPK model.  $\theta'$  can be subdivided in three groups: physiological parameters, butadiene-specific parameters and isoflurane partition coefficients. We randomly picked 100 subjects out of the 133 studied by Mezzetti *et al.* (2003) and used the mode of the joint posterior parameter distribution ( $\Theta_0$ , vector of parameter estimates with the highest posterior probability) obtained by Mezzetti *et al.*, to set values for their physiological characteristics and butadiene-specific parameters. For each subject, individual characteristics (age, sex ...) were kept the same as in the previous study. Using posterior parameter values and measured characteristics of actual subjects of a previous study allowed us to retain correlations between model parameters.

Isoflurane partition coefficients ( $PC'_{a/b}$ ,  $PC'_{pp}$  and  $PC'_{wp}$ ) had not been measured (or estimated) for those subjects and we sampled them from a lognormal distribution with mean 1.4, 3.4, and 1.8 and standard deviation 1.2 (in log scale) (Vinegar *et al.*, 1998). Fat partition coefficient of butadiene and isoflurane was respectively set to 22 (Mezzetti *et al.*, 2003) and 52 (Vinegar *et al.*, 1998).

### ***Simulated experimental designs (Step B, symbol X in Figure 1)***

Butadiene exposures were defined by combining four inhaled concentrations (2, 1, 0.8 and 0.5 ppm) and three exposure times (20, 10 and 5 minutes) yielding twelve concentration - time designs. For each of these designs, ten timed breath samples (5 during exposure and 5 afterward) were simulated for each subject. Measurement times depended on exposure duration. For all 20 minutes exposures, exhaled breath samples were simulated at times 2, 5, 10, 15, 19, 21, 22, 28, 38 and 57 minutes, as proposed by Bois *et al.* (1999b). For 10 minutes exposures, time scale was uniformly reduced by a factor 2 and samples were simulated at times 1, 2, 5, 7, 9, 11, 12, 14, 19 and 28 minutes. For five minutes exposures, time scale was shrunk again but the short exposure duration would not allow taking five samples during exposure, and only nine samples (1, 2, 3, 4, 6, 7, 8, 9 and 14 minutes) were simulated.

Only one experimental design was defined for isoflurane exposure. Each subject was supposed to be exposed to 20 ppm of isoflurane for 15 minutes. Samples of exhaled breath were simulated at 2, 4, 8, 12, 14, 16, 17, 21, 28 and 43 minutes.

### ***Data simulation (Step B, symbol Y in Figure 1)***

Having defined model parameters values and experimental designs, we simulated concentrations of butadiene and isoflurane in exhaled breath at the specified times using the three-compartment PBPK model. Random variations were added to the simulated concentrations to mimic measurement errors that would affect actual experimental measurements. We assumed that observed concentrations would be normally

distributed around their actual values with a standard deviation,  $SD$ , corresponding to the sum of a fixed error plus an error proportional to the measured concentrations (Bois *et al.*, 1994; Kashuba *et al.*, 1996):

$$SD = \delta \times C_{exh}(t) + (1 - \delta) \times LQ \quad (11)$$

When  $C_{exh}(t)$  is at the limit of quantification ( $LQ$ ),  $SD$  is equal to the measured value. When  $C_{exh}(t)$  is largely superior to  $LQ$ , the variation coefficient of the measurement error is equals to  $\delta$ . Data were simulated with  $\delta$  set to 5%, and  $LQ$  was set to 0.006 ppm for butadiene (Smith *et al.*, 2001) and 0.023 ppm for isoflurane (US Department of Labor Occupational Safety and Health Administration, 1994).

This methodology was used to simulate isoflurane data, and butadiene data for 2 ppm exposure concentrations. To facilitate scenario comparison, a variance reduction technique was used (McGeoch, 1992), and data for the other butadiene exposure scenarios were obtained by apportioning the 2 ppm data to exposure concentrations. For example, to obtain concentrations for 1 ppm scenarios, simulated concentrations were divided by 2.

### ***Bayesian calibration (Step C in Figure 1)***

The three-compartment model was fitted to the simulated data (Y) with Bayesian techniques (Bernardo and Smith, 1994; Gelman *et al.*, 1995). The output obtained was a sample ( $\Theta$ ) of parameter values from their joint posterior distribution. From Bayes' theorem, the joint posterior distribution of the parameters to estimate is proportional to the data likelihood multiplied by prior distributions of the parameters. The latter summarizes what is known about parameter values before observing the experimental data (for example, through the literature).

Since our toxicokinetic model has a physiological basis, we had good prior information on parameter values ( $P$  and  $P_{iso}$ ). Each parameter was assumed to have a truncated log-normal density. Geometric means were set to values found in the literature, for physiological characteristics and butadiene (Mezzetti *et al.*, 2003) and isoflurane (Vinegar *et al.*, 1998) parameters. Variances were chosen to correspond to a 50% coefficient of variation, except for butadiene  $k_{met}$  for which the variance was chosen to correspond to a 100% coefficient of variation. We usually expect a priori a larger variability on metabolic parameters. Table 1 sums up the prior distributions used.

**Table 1:** Bounded prior uniform distributions used for the model parameters.

Parameter	Symbol	Prior mean and range
Relative volumes		
Well-perfused tissues	$sc\_V_{wp}$	0.10 [0.01; 0.40]
Fat	$sc\_V_{fat}$	0.20 [0.05; 0.50]
Relative flows		
Poorly perfused tissues	$sc\_F_{pp}$	0.19 [0.08; 0.40]
Fat	$sc\_F_{fat}$	0.05 [0.03; 0.35]
Pulmonary characteristics		
Ventilation rate	$F_{pul}$	7.00 [2.93; 13.4]
Dilution factor	$f_{DS}$	0.40 [0.37; 0.73]
Ventilation perfusion rate	$VPR$	1.14 [0.90; 1.50]
Butadiene partition coefficients		
Blood over air	$PC_{a/b}$	1.50 [0.70; 2.55]
Poorly perfused tissues over blood	$PC_{pp}$	0.80 [0.50; 2.00]
Well-perfused tissues over blood	$PC_{wp}$	1.80 [0.50; 0.40]
Fat over blood	$PC_{fat}$	22 <sup>a</sup>
Butadiene metabolic rate	$k_{met}$	0.30 [0.01; 0.60]
Isoflurane partition coefficients		
Blood over air	$PC'_{a/b}$	1.40 [0.10; 4.00]
Poorly perfused tissues over blood	$PC'_{pp}$	3.40 [0.30; 7.00]
Well-perfused tissues over blood	$PC'_{wp}$	1.8 [0.20; 4.00]
Fat over blood	$PC'_{fat}$	52 <sup>a</sup>

<sup>a</sup> The parameter was set to that value in all simulations.

The likelihood term was given by measurement error models. For each subject, several measurements, affected by uncertainty, were assumed to be available: concentrations of butadiene and isoflurane in exhaled air ( $C_{exh}$ ), butadiene and isoflurane blood-air partition coefficients ( $PC_{a/b}$  and  $PC'_{a/b}$ ), pulmonary ventilation rate ( $F_{pul}$ ), the dilution factor ( $f_{DS}$ ), and the relative fat volume ( $sc\_V_{fat}$ ). We assumed that concentrations in exhaled air had a normal distribution with mean the actual concentrations and standard deviation given by eq (11), the “true” model:

$$\overline{C_{exh}}(t) \sim N(C_{exh}(t), SD) \quad (12)$$

The notation  $\overline{C_{exh}}$  corresponds to the measured butadiene concentrations. The bar is used for measurements affected by errors.

The measured individual characteristics were assumed to be lognormally distributed around their actual values with different standard deviations. The statistical models are:

$$\overline{f_{DS}} \sim LN(f_{DS}, \sigma_1) \quad (13)$$

$$\overline{sc\_V_{fat}} \sim LN(sc\_V_{fat}, \sigma_2) \quad (14)$$

$$\overline{F_{pul}} \sim LN(F_{pul}, \sigma_3) \quad (15)$$

$$\overline{PC_{a/b}} \sim LN(PC_{a/b}, \sigma_4) \quad (16)$$

$$\overline{PC'_{a/b}} \sim LN(PC'_{a/b}, \sigma_4) \quad (17)$$

The standard deviations  $\sigma_1$ ,  $\sigma_2$  and  $\sigma_3$  were set to correspond to a coefficient of variation to 10%. For  $\sigma_4$ , replicate determinations had shown a coefficient of variation of 17% (Bois *et al.*, 1999b). These statistical models were used for the calibration to butadiene and isoflurane data. When only butadiene data were used, the likelihood component for isoflurane blood-air partition coefficient ( $PC'_{a/b}$ ) was omitted.

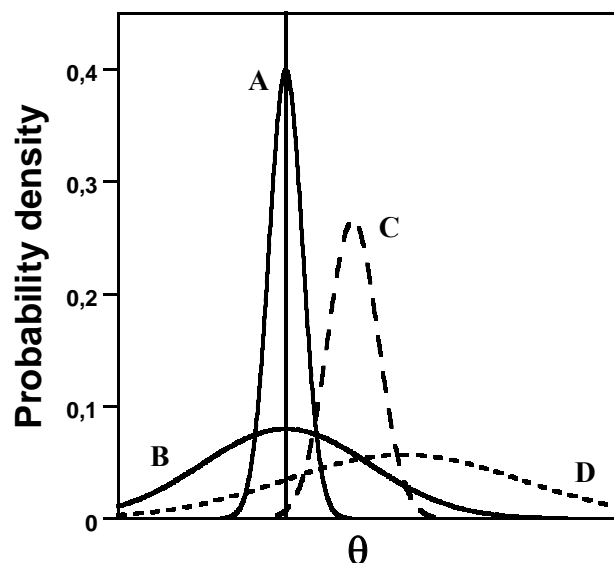
All model calibrations and simulations in Steps A and C were performed using the MCSim software (Bois and Maszle, 1997b). For each analysis in Step C, three independent MCMC chains were run until convergence was reached. Only one in 10 of the 10,000 last iterations was recorded and used to produce the results presented below. For each parameter, the convergence criterion  $\hat{R}$  (Gelman *et al.*, 1995) was used to check convergence to a stable posterior distribution. At perfect convergence, all the  $\hat{R}$  should be equal to 1. Convergence was very likely and was assumed to be achieved when they were all below 1.2.

### ***Comparison of experimental designs***

To quantify butadiene metabolism for all the designs evaluated in Step C, two quantities were estimated for each subject: the metabolic rate constant ( $k_{met}$ , a model parameter) and the fraction of butadiene metabolized during and up to 24 hours after exposure ( $F_{met}$ ). The latter is given by the ratio between the amount metabolized (eq. 2) and the amount inhaled during exposure, which was obtained by multiplying the alveolar ventilation rate by the exposure time and concentration. For each exposure scenario, we then obtain a distribution for  $k_{met}$  and  $F_{met}$ .

To compare distributions obtained with different designs, three criteria were used. First, relative bias and precision (inverse of variance) were evaluated. However, it can be difficult to choose between two distributions (*i.e.*, designs) when both bias and precision affect them. For example, how to choose between a flat unbiased distribution and a precise biased distribution (cases C and D of Figure 3)? To solve this problem, we have chosen to rank designs according to the probability density of the true value under the probability distribution of the parameter of interest ( $P_{true}$  criterion). Figure 3 presents four possible cases. The  $P_{true}$  values are respectively 0.399, 0.080, 0.008 and 0.035 for cases A, B, C and D. Consequently, they can be sorted in the order A, B, D and C. To compute true values densities, we approximated the distribution of each  $k_{met}$  sample by a lognormal distribution, and the  $F_{met}$  one by a normal distribution. Parameters were the empirical mean and standard deviation of our MCMC generated samples. The adequacy of the lognormal and normal approximation was verified with Q-Q plots.

All the simulated experimental designs were first evaluated on a small population of 10 subjects and their results were compared to those obtained for the "reference" exposure to 2 ppm of butadiene during 20 minutes. However it is not possible to finely compare results obtained for scenarios with different exposure times, since only one dataset was considered for 10 subjects and exposure scenario. After the first screening, the most promising experimental designs, combining exposure to butadiene and isoflurane, were then tested on a simulated human population of 100 subjects.



**Figure 3:** Examples of four experimental designs leading to four posterior distribution estimates of a parameter  $\theta$ , with different bias and precision. The vertical line crosses the abscissa axis at the "true" value of the parameter of interest. The four distributions can be ranked in order of the probability density of the true value under them ( $P_{true}$ ).  $P_{true}$  values are respectively 0.399, 0.080, 0.008 and 0.035 for cases A, B, C, and D. Consequently, designs can be sorted in the order A, B, D, and C.

## RESULTS

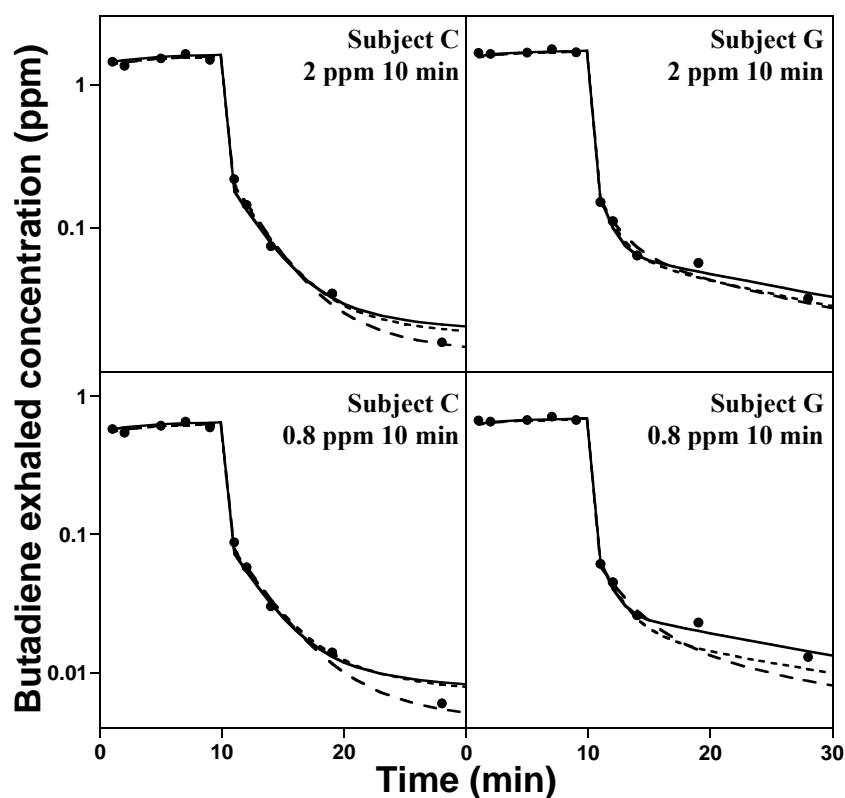
### *Analysis of the experimental designs with 10 subjects*

For each subject, the simulated butadiene data were analyzed either alone or together with the isoflurane data. In each case, 3000 posterior parameter vectors samples were obtained. For the figures presented in this section, the most illustrative subjects were selected.

Figure 4 shows the fit of the model for two subjects and two exposure scenarios (2 ppm or 0.8 ppm butadiene, for 10 minutes). The simulated butadiene data and the true time-course of the butadiene concentration are also represented. To obtain these fits, butadiene concentrations in exhaled air were predicted in several time points with the 3000 posterior parameter vectors, and predictions were averaged in each point. Similar fits were found for all subjects. Figure 4 shows that the estimation of the average time-course of butadiene concentration is improved when isoflurane data are analyzed at the same time. Such an improvement varies from one subject to another. Isoflurane data inclusion can improve significantly the predictions of the butadiene concentration in exhaled air (e.g., Subject C), or more marginally in other cases (e.g., Subject G).

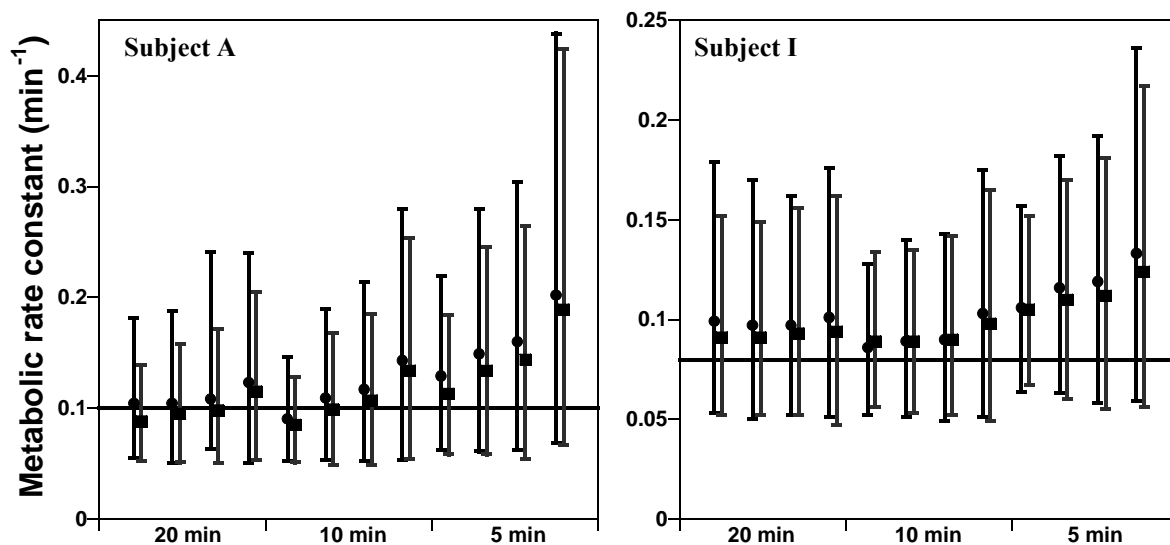
For the twelve butadiene exposure scenarios, we evaluated the influence of the isoflurane data on the estimation of  $k_{met}$  and  $F_{met}$ . On the 10 subjects sample,  $k_{met}$  estimates' relative bias is reduced by 10% to 24%,

precision increases by 7% to 47%, and the probability density of the true value of under the posterior distribution ( $P_{\text{true}}$ ) increases by 10% to 34%. Figure 5 presents the posterior distributions of  $k_{\text{met}}$  for two subjects and for all exposure scenarios with and without isoflurane exposure. The co-analysis of the isoflurane data systematically shortens the 95% confidence interval of  $k_{\text{met}}$ . For a given exposure time, Figure 5 shows that, when butadiene exposure concentration decreases, the 95% confidence intervals of  $k_{\text{met}}$  become wider and the estimated mean becomes larger.



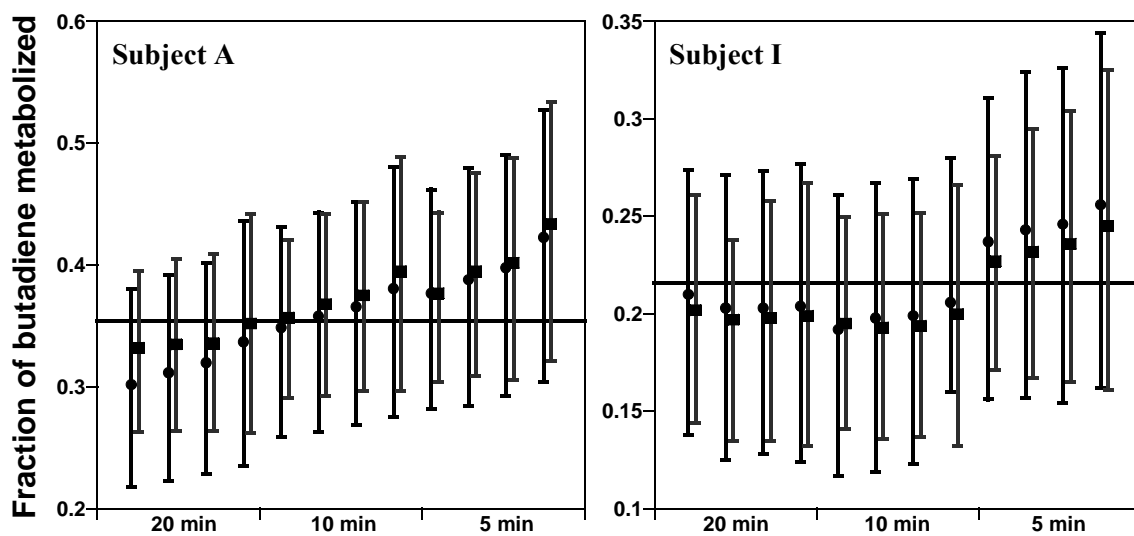
**Figure 4:** Fits of the three-compartment model to butadiene concentrations in exhaled air for subjects C and G, and for two exposure scenarios (2 ppm or 0.8 ppm for 10 minutes). Simulated butadiene data (black dots) were analyzed either alone (resulting fit shown as long dash lines) or together with the isoflurane data (resulting fit shown as short dash lines). The solid line corresponds to the "true" concentrations of butadiene in exhaled air. Fits are point averages of predicted butadiene concentrations (obtained from 3000 posterior parameter vectors).



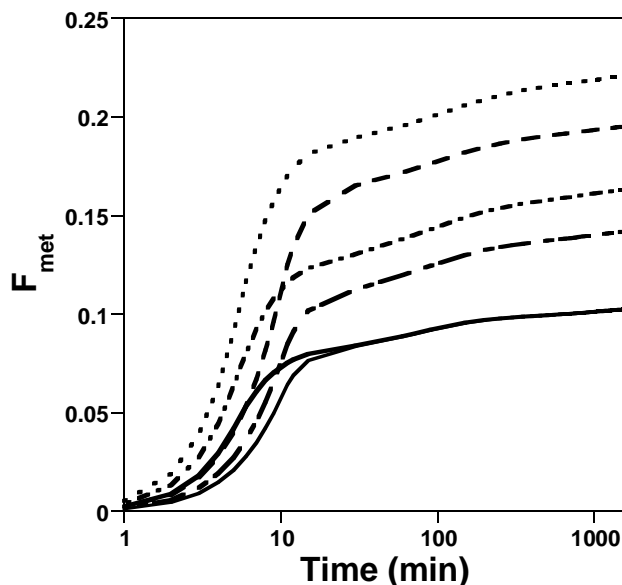


**Figure 5:** Posterior distribution of the butadiene metabolic rate constant,  $k_{met}$ , for two subjects, and twelve butadiene exposure scenarios. For each scenario, the butadiene data were analyzed alone (resulting mean is the black circle) or together with the isoflurane data (mean is the black square). The limits of the bars are the 2.5<sup>th</sup> and the 97.5<sup>th</sup> percentile respectively. The horizontal line corresponds to the true value of  $k_{met}$ . On each panel, the scenarios are rank by exposure time (20, 10 and 5 minutes), and for each time by exposure concentration (2, 1, 0.8 and 0.5 ppm).

For  $F_{met}$ , 24 hours after exposure, the co-analysis of the isoflurane data for the 10 subjects reduces relative bias by 25% to 34%, increases precision by 14% to 51%, and  $P_{true}$  value by 18% to 45%. For two subjects, the posterior predictive distributions of  $F_{met}$  estimates are presented Figure 6 for all exposure scenarios with and without isoflurane exposure. As for  $k_{met}$ , the co-analysis of the isoflurane data increases the precision, and the lower the exposure butadiene concentration, the wider the 95% confidence interval. Figure 7 presents the evolution of  $F_{met}$  for one subject. Two butadiene exposure scenarios are presented (1 ppm for 10 minutes, and 2 ppm for 5 minutes). The time-courses of  $F_{met}$  were produced with and without the analysis of the isoflurane data. These curves were obtained by calculating the  $F_{met}$  values for each of the 3000 posterior parameter vectors at 50 time points. For the selected subject, the predictions of the  $F_{met}$  value are systematically overestimated when based on butadiene data alone. Co-analysis of the isoflurane data reduces significantly this bias.



**Figure 6:** Distributions of the estimated fraction of butadiene metabolized 24 hours after exposure,  $F_{met}$ , for two subjects, and twelve butadiene exposure scenarios. For each scenario, the butadiene data were analyzed alone (resulting mean is the black circle) or together with the isoflurane data (mean is the black square). The limits of the bars are the 2.5<sup>th</sup> and the 97.5<sup>th</sup> percentile respectively. The horizontal line corresponds to the true value  $F_{met}$ . On each panel, the scenarios are rank by exposure time (20, 10 and 5 minutes), and for each time by exposure concentration (2, 1, 0.8 and 0.5 ppm).

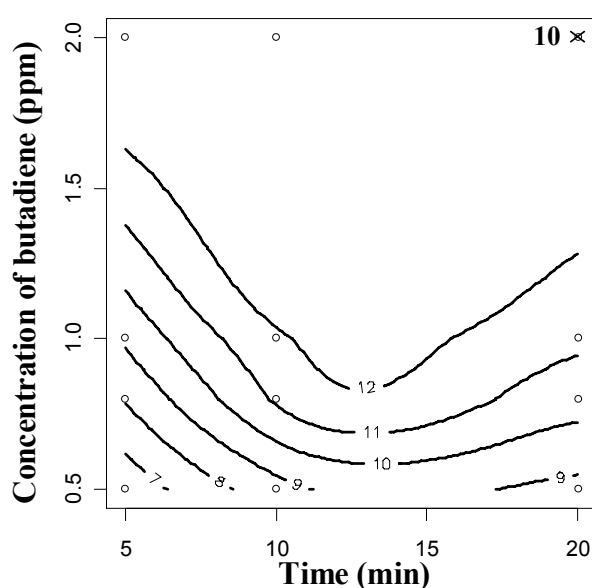


**Figure 7:** Time evolution of the fraction of butadiene metabolized up to 24 hours after exposure,  $F_{met}$ , for Subject C. The time-courses of  $F_{met}$  were produced with and without the analysis of the isoflurane data. The solid lines correspond to the "true" fraction of butadiene metabolized, thin line for the exposure to 1 ppm of butadiene for 10 minutes and thick line for the exposure to 2 ppm of butadiene for 5 minutes. The long dash line corresponds to the predictive curves of  $F_{met}$  for an exposure to 1 ppm of butadiene for 10 minutes, the dotted line to that for an exposure to 2 ppm of butadiene for 5 minutes. The predictive curves obtained for analysis of butadiene data together with isoflurane data are represented by symbols —, and — for 1 ppm of butadiene for 10 minutes and 2 ppm of butadiene for 5 minutes, respectively. Predictive curves were obtained by calculating the  $F_{met}$  values for each of the 3000 posterior parameter vectors at 50 time points.

For the  $k_{met}$  and  $F_{met}$  estimates of the 10 subjects, considering the isoflurane data is always beneficial for the  $P_{true}$  value, in each exposure scenario tested. The benefits of the isoflurane co-analysis on the  $F_{met}$  estimates are superior to those obtained for  $k_{met}$ . This can be explained by the fact that information needed for the  $k_{met}$  estimation is largely contained in the butadiene data, whereas  $F_{met}$  is more influenced by individual physiological parameter (pulmonary flow, for example), which are informed by both isoflurane and butadiene data.

On the basis of these first results, we wanted to select new exposure protocols to evaluate butadiene metabolism on a larger population, with both butadiene and isoflurane exposures. These exposures were to lead to estimates having similar  $P_{true}$  values (a compromise between bias and precision) to those obtained with our "reference exposure scenario" (2 ppm of butadiene for 20 minutes). The  $P_{true}$  values for  $k_{met}$  and  $F_{met}$  of the butadiene reference scenario are respectively 10.3 and 5.4. Figure 8 shows a contour plot of  $P_{true}$  for  $k_{met}$  as a function of butadiene exposure concentration and duration. Figure 9 gives the contours of  $P_{true}$  for

$F_{met}$ . Contour plots were performed with the *R* Software. On both figures, contours were generated with the *loess* function, a nonparametric smoothing method. When exposure concentration and duration are reduced,  $P_{true}$  decreases. The contours were obtained by analysis of 10 subjects' datasets. We selected exposure scenarios that gave equal performance as the reference scenario. On the figures, the contour of interest (10 for  $k_{met}$  and 5 for  $F_{met}$ ) can be followed to select equivalent exposure scenarios. Given Figure 8 and 9, we then retain three scenarios combining butadiene and isoflurane exposures; scenario I: 0.8 ppm butadiene for 20 minutes, scenario II: 1 ppm for 10 minutes, and scenario III: 2 ppm for 5 minutes. These scenarios lead respectively to a total butadiene exposure dose of 16 ppm x min, 10 ppm x min and 10 ppm x min.



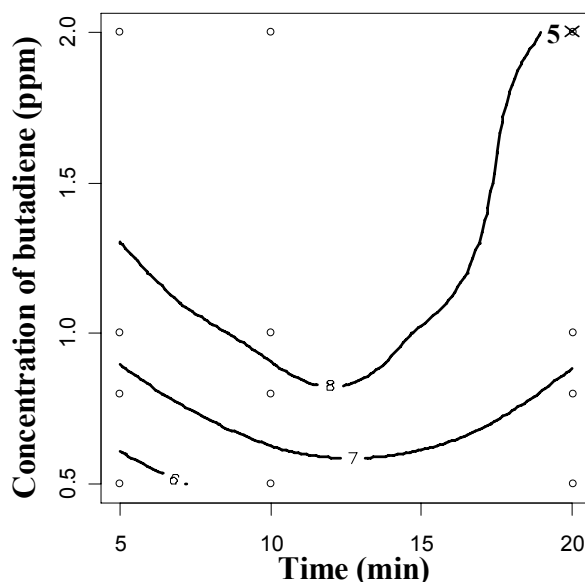
**Figure 8:** Contour plot of the probability density of the true  $k_{met}$  value under its estimated distribution ( $P_{true}$ ), as a function of butadiene exposure concentration and duration. Contours give the  $P_{true}$  values obtained by co-analysis of butadiene and isoflurane data. The numerical value of "10" next to the point (20 minutes and 2 ppm) is  $P_{true}$  for the "reference" scenario, without isoflurane exposure.

### Testing selected experimental designs on 100 subjects

Scenarios I, II and III, and the "reference" scenario (2 ppm of butadiene for 20 minutes) were tested on a population of 100 subjects to confirm their usefulness.  $P_{true}$  values for  $k_{met}$  were very close to those found on the smaller population (reference scenario: 9.9; scenario I: 9.9; scenario II: 10.0; scenario III: 10.1).

The total butadiene dose (16 ppm x min) for scenario I is 1.6 times superior to that for scenarios II and III (10 ppm x min). Therefore, in this example, we would choose to expose subjects to 20 ppm of isoflurane for 15 minutes and to 1 ppm of butadiene for 10 minutes (or to 2 ppm of butadiene for 5 minutes) in place of

a butadiene-only exposure to 2 ppm for 20 minutes. This would lead to a factor 4 reduction in the total dose of butadiene received by the subjects while maintaining the overall information content of the experiments.



**Figure 9:** Contour plot of the probability density of the true  $F_{met}$  value under its estimated distribution ( $P_{true}$ ), as a function of butadiene exposure concentration and duration. Contours give the  $P_{true}$  values obtained by co-analysis of butadiene and isoflurane data. The numerical value of "5" next to the point (20 minutes and 2 ppm) is  $P_{true}$  for the "reference" scenario, without isoflurane exposure.

## DISCUSSION

Controlled human exposures might be acceptable only if four conditions are fulfilled (NRC, 2004): the toxicological question cannot be answered with animal or *in vitro* studies, the benefits outweigh any anticipated risks to exposed subjects, subjects will experience no adverse effects, all ethical standards and procedures are observed. There is therefore a need for design methods leading to low as possible risks for the exposed volunteers. In this work, we propose to reduce the dose of toxic substance received by human subjects during controlled laboratory exposures, and to compensate the loss of information by exposing the same subjects to a non-toxic probe. We are not aware of any published work involving a similar use of chemical probe and toxicant exposures.

We applied our methodology to develop 1,3-butadiene exposure designs for humans, and use isoflurane, an anaesthetic gas, as the non-toxic probe. Analyzing both butadiene and isoflurane experimental data reduced uncertainty in parameter estimates, even for butadiene-specific ones (such as the metabolic rate constant,  $k_{met}$ , or the fraction of butadiene metabolized,  $F_{met}$ ). Alternatively, with a co-analysis of probe data,

the total butadiene exposure dose could be reduced by a factor 4, while maintaining information content (as judged by a compromise between bias and precision). Obviously, having to test two chemicals adds a burden in terms of experimental designs, analytical chemistry, and data analysis. But the benefits for the study subjects are likely to outweigh those considerations.

The two substances used in this study (1,3-butadiene and isoflurane) are both gases administered by inhalation, and have similar biodistributions. Therefore the same PBPK model (*i.e.*, analytical form of the model) was used to describe their ADME processes (the only difference was the absence of isoflurane metabolism). Probe and toxicant kinetics had therefore common physiological parameters. The co-analysis of the probe data directly informed the individual physiological parameters, leading to reduce uncertainty in all model parameter estimates. It would be possible to describe the kinetics of the two substances with very different toxicokinetic models. However, those models should still have some parameters in common. For example, it is preferable to couple substances having similar distribution and tissue affinity (for example, two lipophilic or hydrophilic substances). The choice of substances could actually be guided by a sensitivity analysis of the parameters controlling the toxicant kinetics.

We have assumed that the two chemical substances were not administered jointly. However, experimentalists may prefer to expose the subjects to the two substances at the same time to simplify procedures. In that case, possible chemical interactions between the two substances should be avoided, or modeled if they occur. If experimentalists decide not to expose subjects to the substances at the same time, the intra-subject variation should be considered, and taken into account if it is non-negligible.

Our findings are strongly conditioned by the heteroscedastic error model (eq. 11) we used. Simpler models (*e.g.*, constant relative error) would not lead to such conclusions, because those models implicitly assume that information content is preserved when exposure dose is lowered. However it is well-known of experimentalists that limits of quantification do exist and that precision degrades in its neighborhood. We were careful also to use realistic values for the error model values ( $LQ$  and  $\delta$ ). For the predictive curves we obtain (Figure 4), the relative error is higher at low concentrations, than at the plateau. This is consistent with actual observations (Mezzetti *et al.*, 2003). We are therefore confident that the error model used, and hence our simulations, provide a realistic context for scenario evaluations.

We first evaluated our new exposure protocols on 10 subjects, and then on 100 subjects. In the two cases, only one set of butadiene and isoflurane data was simulated for each subject and exposure scenario. In fact, that is valid when inter-individual variability is higher than intra-individual variability. This condition has been verified on 11 human subjects exposed twice to butadiene in similar experimental conditions as those used here (data not shown). That analysis has showed that inter-individual variability of the metabolic rate constant ( $k_{met}$ ) was on average 7 times higher than its intra-individual variability. Hence, we have chosen, for butadiene, to simulate the same experiment for a large number of subjects, rather than replicate the same experiment on few subjects. A similar evaluation of the Monte Carlo sampling should be performed if other substances were to be investigated.

The methodology we proposed in this paper could be improved at different levels. To define our experimental designs, we used a sampling schedule formally optimized in a previous study (Bois *et al.*, 1999b) for the reference butadiene exposure (2 ppm during 20 minutes). That optimization specified the number and timing of the exhaled air samples collected, in order to obtain a precise estimate of butadiene  $k_{met}$ . Here, time measurements for 10 minutes and 5 minutes designs were simply adapted from that original sampling schedule, to reduce as much as possible the difference between designs and facilitate their comparisons. Yet, the resulting timings may not be optimal for 10 minutes and 5 minutes designs. Nevertheless, if the methodology developed by Bois *et al.* (1999b) had been applied to 10 and 5 minutes exposure scenarios, we suspect that the resulting parameter estimates would be better (in terms of bias and precision) than those obtained with our "non-optimized" sampling schedules. That could only lead to reinforce the choice of the 10 and 5 minutes exposure scenarios, in order to lower butadiene exposure dose.

The model calibration was here performed independently for each subject. In toxicokinetic or pharmacokinetic studies involving groups of individuals, population models are tools of choice to analyze all individual datasets together and take into account several levels of variability and uncertainty (Gelman *et al.*, 1996b; Sheiner and Ludden, 1992). In a population framework, information on each subject is reinforced by the information contained in the other subjects' data, and the estimation process is overall improved (Bois, 2001). For our study, analyzing the 10 or 100 subjects' datasets in Step C with such an approach would probably improve somewhat the parameters' estimate. Again, this would reinforce our choice of lower exposure scenarios, without changing our overall conclusions. Our primary aim was to show that analyzing probe data could help to lower the toxicant exposure dose, and we did not want to complicate the analysis with a population toxicokinetic model. Adding such a model would be a logical following of this design study for the analysis of actual experiments.

On a similar note, we have noticed that there are differences in the benefit of the co-analysis of the probe data between subjects. We did not perform further analysis to understand which factors drive these differences. If such factors exist, it could be useful to identify them and to select subjects for future exposure studies in order to maximize the benefits of the co-analysis of the chemical probe.

Overall, our method provides an easy way to lower the toxicant exposure dose received by subjects (in particularly, humans) during controlled exposures by exposing the same subjects to a non-toxic substance. Our findings are based on a simulation study, and should obviously be confirmed with actual data.

## REFERENCES

- Andersen, M. E. (1995). Development of physiologically based pharmacokinetic and physiologically based pharmacodynamic models for applications in toxicology and risk assessment. *Toxicol. Lett.* 79, 35-44.
- Bailer, A. J., and Dankovic, D. A. (1997). An introduction to the use of physiologically based pharmacokinetic models in risk assessment. *Statistical Methods in Medical Research* 6, 341-358.
- Bernardo, J. M., and Smith, A. F. M. (1994). *Bayesian Theory*. Wiley, New York.

- Bernillon, P., and Bois, F. Y. (2000). Statistical issues in toxicokinetic modeling: A Bayesian perspective. *Environmental Health Perspectives* 108, 883-893.
- Bois, F. Y. (2001). Applications of population approaches in toxicology. *Toxicol. Lett.* 120, 385-394.
- Bois, F. Y., and Maszle, D. (1997). MCSim: a simulation program. *Journal of Statistical Software* 2(9), <http://www.stat.ucla.edu/journals/jss/v02/i09> (software available online at [www.toxi.ineris.fr](http://www.toxi.ineris.fr)).
- Bois, F. Y., Smith, T. J., Gelman, A., Chang, H. Y., and Smith, A. E. (1999). Optimal design for a study of butadiene toxicokinetics in humans. *Toxicol. Sci.* 49, 213-224.
- Bois, F. Y., Tozer, T. N., Hauck, W. W., Chen, M. L., Patnaik, R., and Williams, R. L. (1994). Bioequivalence: performance of several measures of extent of absorption. *Pharm. Res.* 11, 715-22.
- Carpenter, R. L., Eger, E. I., 2nd, Johnson, B. H., Unadkat, J. D., and Sheiner, L. B. (1986). The extent of metabolism of inhaled anesthetics in humans. *Anesthesiology* 65, 201-5.
- Deurenberg, P., Weststrate, J. A., and Seidell, J. C. (1991). Body mass index as a measure of body fatness: age- and sex-specific prediction formulas. *Br J Nutr* 65, 105-14.
- Gelman, A., Bois, F. Y., and Jiang, J. (1996). Physiological pharmacokinetic analysis using population modeling and informative prior distributions. *J. Am. Stat. Assoc.* 91, 1400-1412.
- Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (1995). *Bayesian Data Analysis*. Chapman & Hall, London.
- Haerberle, H. A., Wahl, H. G., Jakubetz, H., Krause, H., Schmidt, R., Schurig, V., and Dieterich, H. J. (2002). Accumulation of S(+) enantiomer in human beings after general anaesthesia with isoflurane racemate. *Eur J Anaesthesiol* 19, 641-6.
- Himmelstein, M. W., Turner, M. J., Asgharian, B., and Bond, J. A. (1996). Metabolism of 1,3-butadiene: inhalation pharmacokinetics and tissue dosimetry of butadiene epoxides in rats and mice. *Toxicology* 113, 306-9.
- International Agency for Research on Cancer (1992). *Occupational Exposures to Mists and Vapours from Strong Inorganic Acids and Other Industrial Chemicals*. International Agency for Research on Cancer, Lyon, France.
- Kashuba, A. D., Ballow, C. H., and Forrest, A. (1996). Development and evaluation of a Bayesian pharmacokinetic estimator and optimal, sparse sampling strategies for ceftazidime. *Antimicrob Agents Chemother* 40, 1860-5.
- McGeoch, C. (1992). Analyzing algorithms by simulation: variance reduction techniques and simulation speedups. *ACM Computing Surveys* 24, 195-212.
- Mezzetti, M., Ibrahim, J. G., Bois, F. Y., Ryan, L. M., Ngo, L., and Smith, T. J. (2003). A Bayesian compartmental model for the evaluation of 1,3-butadiene metabolism. *Journal of the Royal Statistical Society Series C-Applied Statistics* 52, 291-305.
- National Research Council - Committee on the Use of Third Party Toxicity Research with Human Research Participants Science Technology and Law Program (NRC) (2004). *Intentional Human Dosing Studies for EPA Regulatory Purposes: Scientific and Ethical Issues*. The National Academies Press.



Rowland, M., and Tozer, T. N. (1989). *Clinical Pharmacokinetics: Concepts and Applications*. Lea & Febiger, Philadelphia.

Sheiner, L. B., and Ludden, T. M. (1992). Population pharmacokinetics/dynamics. *Annual Review of Pharmacology and Toxicology* 32, 185-209.

Smith, T., Lin, Y.-S., Mezzetti, L., Bois, F. Y., Kelsey, K., and Ibrahim, J. (2001). Genetic and dietary factors affecting human metabolism of 1,3-butadiene. *Chemico-Biological Interactions* 135-136, 407-428.

US Department of Labor Occupational Safety and Health Administration (1994). Sampling and Analytical Methods, Enflurane, Halothane, Isoflurane, Method number: 103. <http://www.osha.gov/dts/sltc/methods/index.html>.

Vinegar, A., Jepson, G. W., and Overton, J. H. (1998). PBPK modeling of short-term (0 to 5 min) human inhalation exposures to halogenated hydrocarbons. *Inhal. Toxicol.* 10, 411-429.

Watanabe, K., and Bois, F. Y. (1996). Interspecies extrapolation of physiological pharmacokinetic parameter distributions. *Risk Analysis* 16, 741-754.

## **Lumping in pharmacokinetics**

Céline Brochot<sup>1</sup>, János Tóth<sup>2</sup>, and Frédéric Y. Bois<sup>1</sup>

<sup>1</sup>INERIS, Institut National de l'Environnement Industriel et des Risques, Unité de Toxicologie Expérimentale, Parc Alata BP2, 60550 Verneuil En Halatte, France

<sup>2</sup>Department of Mathematical Analysis, Budapest University of Technology and Economics, Hungary

En préparation (soumission prévue à *Journal of Pharmacokinetics and Biopharmaceutics*)

## RESUME

Pour décrire la biodistribution d'une substance chimique, les modèles pharmacocinétiques sont typiquement utilisés. Ces modèles simplifient la complexité biologique en divisant l'organisme en compartiments interconnectés. L'évolution temporelle de la quantité (ou concentration) de substance dans chaque compartiment est décrite par un système d'équations différentielles. La complexité d'un tel système peut augmenter rapidement, si une description précise des processus biologiques est nécessaire. Cependant, la manipulation de systèmes à grande dimension peut s'avérer difficile, notamment pour l'optimisation d'un protocole d'exposition.

Pour surmonter ces difficultés, des méthodes d'agrégation mathématiques ont été développées. Ces méthodes ont pour but de réduire un système d'équations différentielles en agrégeant plusieurs variables entre elles. Typiquement, le modèle agrégé est toujours un système d'équations différentielles, dont les variables sont interprétable en termes des variables du système original. En pratique, le modèle agrégé doit satisfaire quelques restrictions. Par exemple, il peut être nécessaire de ne pas agréger certaines variables d'intérêt. Pour tenir compte de ces restrictions, des méthodes d'agrégation contrainte ont été développées, et sont appliquées avec succès dans l'industrie chimique et pétrolière.

Dans ce travail, nous proposons d'étudier, à l'aide d'exemples pratiques, le potentiel de ces méthodes pour la toxico/pharmacocinétique. Comme application, nous simplifions un modèle à 2 compartiments par des méthodes d'agrégation symbolique. Ensuite, nous proposons de réduire numériquement un modèle physiologique à 6 compartiments pour le 1,3-butadiène avec des méthodes d'agréations contraintes. Les méthodes d'agrégation, présentées ici, peuvent être automatisées facilement, et sont applicables à de nombreux systèmes d'équations différentielles.

## ABSTRACT

Pharmacokinetic models simplify biological complexity by dividing the body into interconnected compartments. The time course of the chemical's amount (or concentration) in each compartment is then expressed as a system of ordinary differential equations. The complexity of the resulting equation system can rapidly increase, if a precise description of the organism is needed. However, computation of high-dimensional models is difficult and lengthy, for example when optimizing an experimental protocol.

To overcome such difficulties, mathematical lumping methods are available. Such methods aim at reducing a differential equation system by aggregating several variables into one. Typically, the lumped model is still a differential equation system, whose variables are interpretable in terms of variables of the original system. In practice, the reduced model is usually required to satisfy some restrictions. For example, it can be necessary to keep unlumped the state variables of interest for prediction. To accommodate such restrictions, constrained lumping methods have also been developed. These lumping methods have been successfully applied in the fields of atmospheric or petroleum chemistry and combustion.

We propose, here, to study, through practical examples, the potential of such methods in toxico/pharmacokinetics. As a tutorial, we first simplify a 2-compartment pharmacokinetic model by symbolic lumping. We then explore the reduction of a 6-compartment physiologically based pharmacokinetic model for 1,3-butadiene with numerical constrained lumping. The lumping methods, presented here, can be easily automated, and are applicable to any kind of differential equation systems.

## INTRODUCTION

Two kinds of pharmacokinetic (PK) models, or toxicokinetic (TK) models for toxic compounds, are typically used to describe the absorption, distribution, metabolism and elimination of chemicals as a function of time: data-based compartmental PK models and physiologically based pharmacokinetic (PBPK) models (Gibaldi and Perrier, 1982). All simplify biological complexity by dividing the body into interconnected compartments. The time evolution of the chemical's amount (or concentration) in each compartment is governed by a system of ordinary differential equations. The dimension of such systems depends on model complexity. Typically, the finer the description of the biodistribution process, the higher the system dimension. Due to their pretense at describing anatomical and physiological structure, PBPK models have typically higher complexity and dimensionality than compartmental PK models. However, difficulties arise when the PK model contains more variables and parameters to be comfortable for mathematical and computational treatment. Indeed parameter estimation, dosage design and optimization are not easy to handle with high-dimensional models.

Lumping techniques aim at reducing model dimensionality and complexity by aggregating several variables into one. The aggregation can concern different species (sometimes, unknown as in combustion process), in the case of concomitant exposure to multiple interacting agents for example (Dennison *et al.*, 2004; Dennison *et al.*, 2003). At first sight, one needs  $n$  PK models to describe the behavior of  $n$  substances. In this case, to simplify modeling and calculations, agents exhibiting the same physico-chemical and biodistribution properties are lumped in a single "species". Only one PK model is therefore needed to model the kinetics of this new species. Another lumping approach, the one of interest in this paper, is the lumping of the model's mathematical structure. Usually, the lumped model is still an ordinary differential equations system with new variables, interpretable in terms of variables of the original system. So far, in pharmacokinetics, only semi-empirical methods have been proposed (Bjorkman, 2003; Nestorov *et al.*, 1998). The latter have proposed a procedure applicable to PBPK models: only tissues/compartments with identical specifications (e.g., similar transfer rate) and occupying identical positions in the system structure can be lumped together. For example, lumped tissues should have similar or close time constants, or should equilibrate very rapidly with each other. Hence, the lumping process depends on parameter values, and should be re-evaluated for each (class of) substance(s).

Alternatively, several mathematical (exact or approximate) lumping methods have been proposed (Li and Rabitz, 1989, 1990). These methods are only based on the mathematical structure of the differential equations system, and are applicable to linear or non-linear system (Li and Rabitz, 1989; Li *et al.*, 1994a; Li *et al.*, 1994b). Such methods have been originally developed in the fields of atmospheric or petroleum chemistry and combustion (Li and Rabitz, 1993; Tomlin *et al.*, 1997; Wei and Kuo, 1969). They have the potential (demonstrated through past applications) to provide effective solutions to the need for a computational approximation of a chemical mechanism (Li and Rabitz, 1993; Verhaar *et al.*, 1997; Wang *et al.*, 1998). In practice, the reduced model is usually required to satisfy some restrictions. For example, it can

be necessary to keep unlumped some state variables which have been experimentally measured (such as blood concentration), or which are of interest for prediction (for example, the quantity metabolized). To accommodate such restrictions, constrained lumping methods have also been developed (Li and Rabitz, 1991a, c, 1993). Of course, the choice of lumping method depends essentially on the objectives of the study and on the model structure.

In this paper, we focus on linear PK models, *i.e.* corresponding to linear differential equation systems. We present mathematical lumping methods applicable to such models, together with examples. First, the fundamentals of unconstrained and constrained lumping for linear differential equations systems are briefly introduced. Then, symbolic lumping is applied to a general two-compartment model. Finally, a PBPK model for 1,3-butadiene biodistribution is treated by numerical constrained lumping.

## METHODS

In this section, we present the constrained and unconstrained lumping methods for linear systems of first order differential equations. For simplicity, some standard mathematical assumptions are omitted. More details can be found in Li and Rabitz (1989; 1991a).

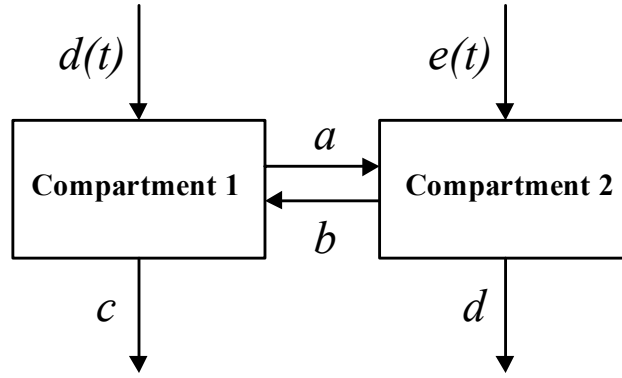
### *PK models and the lumping process*

A general definition of a PK model (of dimension  $n$ ) can be given by

$$\frac{\partial \mathbf{y}(t)}{\partial t} = \mathbf{K}\mathbf{y}(t) + \mathbf{u}(t) \quad (1)$$

where  $\mathbf{y}$  is the vector containing the  $n$  state variables (usually the amount or concentration of the chemical in a compartment),  $\mathbf{u}$  the inflow vector (such as through inhalation) and  $\mathbf{K}$  the matrix of coefficients (of dimension  $n \times n$ ). For data-based compartmental PK models, elements of  $\mathbf{K}$  are usually the transfer rate constants (or linear combinations of them) between the compartments. For PK models, some restrictions on the elements of  $\mathbf{K}$  occur (Jacquez and Simon, 1993):

$$\begin{aligned} K_{ij} &\geq 0 & i \neq j; j = 1 \dots n \\ K_{ii} &\leq -\sum_{j=1}^n K_{ji} & i = 1 \dots n \end{aligned} \quad (2)$$



**Figure 1:** Representation of a general 2-compartment model. Arrows describe the exchange of material.

For example, let us consider a general two-compartment model (Figure 1),

$$\begin{cases} \frac{\partial y_1(t)}{\partial t} = -(a+c)y_1(t) + by_2(t) + d(t) \\ \frac{\partial y_2(t)}{\partial t} = ay_1(t) - (b+f)y_2(t) + e(t) \end{cases} \quad (3)$$

All the rate coefficients are nonnegative real constants. The matrix  $\mathbf{K}$  and the vector  $\mathbf{u}$  are then given by

$$\mathbf{K} = \begin{pmatrix} -(a+c) & b \\ a & -(b+f) \end{pmatrix} \quad \mathbf{u}(t) = \begin{pmatrix} d(t) \\ e(t) \end{pmatrix} \quad (4)$$

To lump the system (eq. 1) into a  $\hat{n}$ -dimensional system ( $\hat{n} \leq n$ ), we introduce some notation. Let  $\hat{\mathbf{y}}$  be the  $\hat{n}$ -dimensional vector containing the state variables of the new lumped system, and  $\hat{\mathbf{u}}$  the new inflow vector. These two vectors are obtained by

$$\hat{\mathbf{y}}(t) = \mathbf{M}\mathbf{y}(t) \quad \text{and} \quad \hat{\mathbf{u}}(t) = \mathbf{M}\mathbf{u}(t) \quad (5)$$

where  $\mathbf{M}$  is a constant matrix (of dimension  $\hat{n} \times n$ ), called "lumping matrix". The system (eq.1) is considered to be lumpable if there exists a matrix  $\hat{\mathbf{K}}$  (dimension  $\hat{n} \times \hat{n}$ ) such that

$$\frac{\partial \hat{\mathbf{y}}(t)}{\partial t} = \hat{\mathbf{K}}\hat{\mathbf{y}}(t) + \hat{\mathbf{u}}(t) \quad (6)$$

Li and Rabitz (1989) proved that such a matrix  $\hat{\mathbf{K}}$  can be obtained as:

$$\hat{\mathbf{K}} = \mathbf{M}\mathbf{K}\overline{\mathbf{M}} \quad (7)$$

where  $\overline{\mathbf{M}}$  is a generalized inverse of  $\mathbf{M}$  (i.e., their matrix multiplication gives the identity matrix of dimension  $\hat{n} \times \hat{n}$ ).

The task is now to construct the lumping matrices  $\mathbf{M}$ . In the following paragraphs, the construction of such lumping matrices is presented in the unconstrained and constrained cases.

### ***Construction of the lumping matrix $\mathbf{M}$***

In the case of linear systems, the construction of the lumping matrix  $\mathbf{M}$  is relatively straightforward (Li and Rabitz, 1989). When the aggregation is exact, the subspace spanned by the rows of  $\mathbf{M}$  should be  $\mathbf{J}^T(\mathbf{y})$ -invariant, with  $\mathbf{J}^T(\mathbf{y})$  the transpose of the Jacobian matrix of  $\mathbf{K}\mathbf{y}$ . This latter can be expanded according to

$$\mathbf{J}^T(\mathbf{y}) = \sum_{k=1}^m a_k(\mathbf{y}) \mathbf{A}_k \quad (8)$$

where  $m$  is less than  $n^2$ , and the constant matrices  $\mathbf{A}_k$  are viewed as a set of basis matrices of  $\mathbf{J}^T(\mathbf{y})$ . For a unimolecular reaction system and in particular for linear PK systems, the Jacobian matrix is  $\mathbf{K}$ ,

$$\mathbf{J}^T(\mathbf{y}) = \mathbf{K}^T \quad (9)$$

Therefore, we have to find the space of dimension  $\hat{n}$  that is invariant to all  $\mathbf{A}_k$  in order to determine  $\mathbf{M}$ .

#### *Unconstrained lumping*

In case of linear PK system, the Jacobian matrix is given by eq. 9. Therefore, the rows of a  $\hat{n}$ -dimensional matrix  $\mathbf{M}$  are composed by the  $\hat{n}$  eigenvectors of the transpose of  $\mathbf{K}$ . Several lumping matrices can be defined for the same system dimension. Indeed,  $n!/(n-\hat{n})!$   $\hat{n}$ -dimensional lumping matrices can be constructed for a  $n$ -dimensional system with distinct eigenvalues. It is also possible to multiply (from the left) such lumping matrices by a nonsingular quadratic matrix without changing their lumping property.

#### *Constrained lumping*

Li and Rabitz (1991a; 1991c) have proposed the direct constrained approximate lumping (DCAL) method to determine the matrix  $\mathbf{M}$  when constraints are imposed on the state variables. The lumping matrix  $\mathbf{M}$  is composed by two sub-matrices, one ( $\mathbf{M}_G$ ) given by the constraints, and the other ( $\mathbf{M}_D$ ) to be determined. The DCAL method determines base matrices  $\mathbf{A}_k$ 's numerically, by using values of  $\mathbf{y}$  in a region of the  $n$ -dimensional composition space of interest. Briefly, the different steps are:

- Transform  $\mathbf{M}_G(\mathbf{A}_k)^T$  ( $i$  from 0 to  $\text{Rank}(\mathbf{A}_k) - 1$ ) into an orthogonal matrix  $\mathbf{Q}(\mathbf{G})_{ki}^T$  by Gram-Schmidt orthogonalization (Bellman, 1960),
- Construct the symmetric matrix  $\mathbf{Y}$  defined by:

$$\mathbf{Y} = \sum_{k=1}^m \sum_{i=0}^{\text{Rank}(\mathbf{A}_k)-1} \mathbf{Q}(\mathbf{G})_{ki}^T \cdot \mathbf{Q}(\mathbf{G})_{ki} \quad (10)$$

- Determine the eigenvalues and eigenvectors of  $\mathbf{Y}$ .

When the eigenvectors of  $\mathbf{Y}$  are arranged according to the decreasing order of the magnitude of their eigenvalues, the first  $\hat{n}$  columns compose the best constrained approximate lumping matrix with dimension  $\hat{n}$ , according to the DCAL method. If  $m$  eigenvalues of  $\mathbf{Y}$  are equal to zero, the first  $n-m$  eigenvectors of  $\mathbf{Y}$  compose an exact lumping matrix  $\mathbf{M}^T$ .



## APPLICATIONS OF MATHEMATICAL LUMPING IN THE PK FIELD

In this section, we apply lumping methods to some PK models. First, a general 2-compartment model is simplified by symbolic unconstrained lumping. A 6-compartment PBPK model for 1,3-butadiene (BD) is then presented and simplified with numerical constrained lumping. The *Mathematica* software was used for all calculations.

### *Symbolic lumping of a general two-compartment model*

Two-compartment models are often used in pharmaco/toxicokinetics. They can describe either the global distribution of a substance (e.g., for 1,3-butadiene, see Bois *et al.*, 1999b) or its local action (e.g., Daldrup *et al.*, 1998). Let us consider the general two-compartment model described in eq. 3. All the rate coefficients are nonnegative real constants with the natural constraint that  $a$  is positive (this does not in fact restrict generality). This system is of the form of the eq. 1, and the matrix  $\mathbf{K}$  and the vector  $\mathbf{u}$  were given in the previous section.  $\lambda_1$  and  $\lambda_2$ , the two eigenvalues of  $\mathbf{K}^T$ , are

$$\begin{aligned}\lambda_1 &= \frac{1}{2} \left( -a - b - c - f + \sqrt{(a + b + c + f)^2 - 4(bc + af + cf)} \right) \\ \lambda_2 &= \frac{1}{2} \left( -a - b - c - f - \sqrt{(a + b + c + f)^2 - 4(bc + af + cf)} \right)\end{aligned}\quad (11)$$

Following the unconstrained lumping method described in the previous section, the lumping matrix  $\mathbf{M}$  is composed by the two eigenvectors of  $\mathbf{K}^T$ :

$$\begin{aligned}s_1 &= \begin{pmatrix} -a + b - c + f + \sqrt{(-a + b - c + f)^2 + 4ab} \\ 2b \end{pmatrix} \\ s_2 &= \begin{pmatrix} -a + b - c + f - \sqrt{(-a + b - c + f)^2 + 4ab} \\ 2b \end{pmatrix}\end{aligned}\quad (12)$$

Two one-dimensional lumping systems are therefore possible:

$$\begin{aligned}\frac{\partial \hat{\mathbf{y}}(t)}{\partial t} &= -\lambda_1 \hat{\mathbf{y}}(t) + s_1 \mathbf{u}(t) \\ \frac{\partial \hat{\mathbf{y}}(t)}{\partial t} &= -\lambda_2 \hat{\mathbf{y}}(t) + s_2 \mathbf{u}(t)\end{aligned}\quad (13)$$

The new variable is therefore a combination of the two original variables, as defined by eq. 5,

$$\hat{\mathbf{y}}(t) = s_i^T \mathbf{y}(t) \quad (14)$$

where  $i$  is equal to 1 or 2, according to the selected lumping scheme. Let's take an example. Consider a 2-compartment model in which exchanges between the two compartments are equal, *i.e.*  $a = b$  (in case of a

diffusion, for example), and rates going out of the compartments are also equal, *i.e.*  $c = f$ . Dividing each element by  $2a$ , the eigenvector  $s_I$  of  $\mathbf{K}^T$  is  $(1; 1)$ . We then obtain the new lumped system

$$\frac{\partial \hat{\mathbf{y}}(t)}{\partial t} = -\lambda_1 \hat{\mathbf{y}}(t) + d(t) + e(t) \quad (15)$$

with

$$\hat{\mathbf{y}}(t) = y_1(t) + y_2(t) \quad (16)$$

We have therefore reduced the 2-compartment model into a 1-compartment one, for which the new state variable is the sum of the two original ones.

### ***Numerical constrained lumping of a whole-body PBPK model for 1,3-butadiene***

In this section, we propose to simplify a 6-compartment whole-body PBPK model for BD. BD is a chemical compound largely used in the production of plastics and synthetic rubber. Studies in rats and mice have demonstrated that its metabolites can cause cancer. For actual applications, we should make sure that the new lumped model is still able to answer the question(s) of interest. For this, we can apply constrained lumping methods. One constraint arise in the BD example: to evaluate the BD toxicity, the total quantity of butadiene metabolized after a given exposure should be quantified by the new model, and therefore should not be lumped. First, the full PBPK model is presented, lumped models are derived and then effects of lumping on the results are studied.

#### *A PBPK model for 1,3-butadiene*

BD kinetics in the human body can be described by a whole-body PBPK model, including six compartments (blood, fat, liver, lungs, slowly perfused tissues – referred here to as “muscle” – and other rapidly perfused tissues – “viscera”) and three sites of metabolism (liver, lungs and viscera) (Kohn and Melnick, 1993). BD enters the body by inhalation. We simplified the Kohn and Melnick’s model by using a first order metabolism term, rather than a Michaelis–Menten term. This linearizes the model, while still keeping its relevance for low concentrations exposures (Smith et al., 2001). The model we used is given by the following system of differential equations (of dimension 7):

$$\left\{ \begin{array}{l}
\frac{\partial Q_B(t)}{\partial t} = -F_C \times \left( \frac{Q_B(t)}{V_B} - \frac{Q_{Lu}(t)}{V_{Lu} \times PC_{Lu}} \right) - \sum_{i \in I} F_i \times \left( \frac{Q_B(t)}{V_B} - \frac{Q_i(t)}{V_i \times PC_i} \right) \\
\frac{\partial Q_{Fa}(t)}{\partial t} = F_{Fa} \times \left( \frac{Q_B(t)}{V_B} - \frac{Q_{Fa}(t)}{V_{Fa} \times PC_{Fa}} \right) \\
\frac{\partial Q_{Mu}(t)}{\partial t} = F_{Mu} \times \left( \frac{Q_B(t)}{V_B} - \frac{Q_{Mu}(t)}{V_{Mu} \times PC_{Mu}} \right) \\
\frac{\partial Q_{Li}(t)}{\partial t} = F_{Li} \times \left( \frac{Q_B(t)}{V_B} - \frac{Q_{Li}(t)}{V_{Li} \times PC_{Li}} \right) - K_{Li} \times Q_{Li}(t) \\
\frac{\partial Q_{Lu}(t)}{\partial t} = F_C \times \left( \frac{Q_B(t)}{V_B} - \frac{Q_{Lu}(t)}{V_{Lu} \times PC_{Lu}} \right) + K_{Vent} \times \left( C_{Inh}(t) - \frac{Q_{Lu}(t)}{V_{Lu} \times PC_{Air}} \right) - K_{Lu} \times Q_{Lu}(t) \\
\frac{\partial Q_{Vi}(t)}{\partial t} = F_{Vi} \times \left( \frac{Q_B(t)}{V_B} - \frac{Q_{Vi}(t)}{V_{Vi} \times PC_{Vi}} \right) - K_{Vi} \times Q_{Vi}(t) \\
\frac{\partial Q_{Met}(t)}{\partial t} = K_{Li} \times Q_{Li}(t) + K_{Lu} \times Q_{Lu}(t) + K_{Vi} \times Q_{Vi}(t)
\end{array} \right. \quad (17)$$

The indices  $B$ ,  $Fa$ ,  $Mu$ ,  $Li$ ,  $Lu$ ,  $Vi$  and  $Met$  stand respectively for blood, fat, muscle, liver, lungs, viscera and metabolites. The symbol  $I$  used in the first equation represents the set  $\{Fa, Mu, Li, Vi\}$ . For each compartment  $i$ ,  $Q_i$ ,  $V_i$  and  $PC_i$  are respectively the BD quantity, the volume, and the tissue over blood partition coefficient.  $F_i$  is the blood flow entering in  $i$  and  $K_i$  the metabolic term.  $PC_{Air}$  is the blood over air partition coefficient,  $K_{Vent}$  the ventilation rate and  $C_{inh}$  the inhaled BD concentration. The exhaled BD concentration is given by the following relationship,

$$C_{exh}(t) = \frac{1}{3} \times C_{inh}(t) + \frac{2}{3} \times \frac{Q_{Lu}(t)}{PC_{Lu} \times V_{Lu}} \quad (18)$$

The experimental conditions simulated (an inhalation exposure of 5 ppm during 2 hours) were defined on the basis of actual experiments (Bois *et al.*, 1999b).

#### A constrained lumping scheme

To apply numerical constrained lumping to the differential equation system (eq. 16), parameters were set to the physiological values referred in (Kohn and Melnick, 1993). Given these values, the matrix of the system coefficients,  $\mathbf{K}$ , is

$$\mathbf{K} = \begin{pmatrix} -244.90 & 0.02 & 1.16 & 10.99 & 160.20 & 30.11 & 0 \\ 4.41 & -0.02 & 0 & 0 & 0 & 0 & 0 \\ 44.20 & 0 & -1.16 & 0 & 0 & 0 & 0 \\ 19.59 & 0 & 0 & -210.72 & 0 & 0 & 0 \\ 122.45 & 0 & 0 & 0 & -610.60 & 0 & 0 \\ 54.24 & 0 & 0 & 0 & 0 & -43.61 & 0 \\ 0 & 0 & 0 & 199.73 & 13.50 & 13.50 & 0 \end{pmatrix} \quad (19)$$

To construct the lumping matrix  $\mathbf{M}$ , we first need to define the sub-matrix  $\mathbf{M}_G$  containing the constraints. Suppose the state variables are sorted in the same order as in the system (eq. 17). Given the constraint defined above (quantity metabolized left unlumped),  $\mathbf{M}_G$  is

$$\mathbf{M}_G = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad (20)$$

To determine  $\mathbf{M}$ , we apply the DCAL method presented in the previous section. Eq. 9 states that the jacobian matrix of the system is  $\mathbf{K}$ . Therefore  $\mathbf{A}_0$  is equal to the transpose of  $\mathbf{K}$ . In order to force  $\mathbf{M}_G$  to be located on the first rows of  $\mathbf{M}$ , we multiply  $\mathbf{M}_G$  by 1000. We then obtain the matrix  $\mathbf{Y}$ ,

$$\mathbf{Y} = \begin{pmatrix} 0.342 & -1.19 \times 10^{-5} & -6.93 \times 10^{-4} & -0.560 & -0.793 & -2.18 \times 10^{-2} & 0 \\ -1.19 \times 10^{-5} & 0 & 0 & 1.56 \times 10^{-5} & 2.83 \times 10^{-5} & 0 & 0 \\ -6.93 \times 10^{-4} & 0 & 0 & 9.09 \times 10^{-4} & 1.65 \times 10^{-3} & 4.28 \times 10^{-5} & 0 \\ -0.560 & 1.56 \times 10^{-5} & 9.09 \times 10^{-4} & 2.725 & 1.131 & 0.106 & 0 \\ -0.793 & 2.83 \times 10^{-5} & 1.65 \times 10^{-3} & 1.131 & 1.926 & 5.41 \times 10^{-2} & 0 \\ -2.18 \times 10^{-2} & 0 & 4.28 \times 10^{-5} & 0.106 & 5.41 \times 10^{-2} & 5.94 \times 10^{-3} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \times 10^6 \end{pmatrix} \quad (21)$$

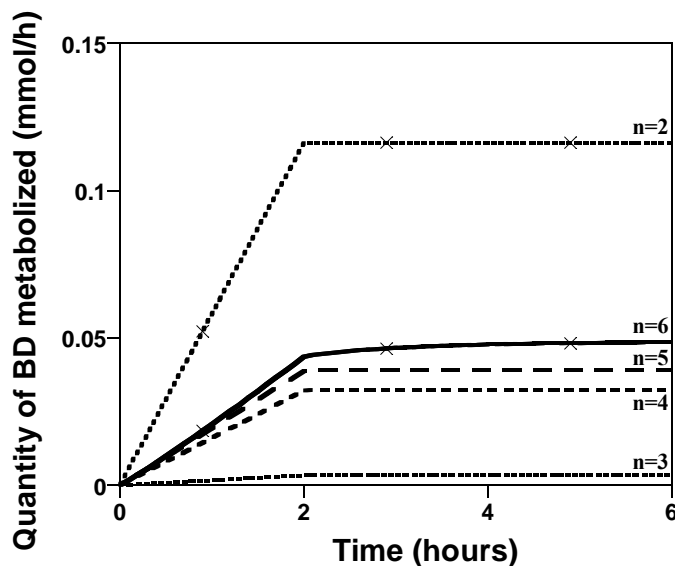
Its eigenvalues are  $(1 \times 10^6, 3.78, 1.21, 0.011, 3.07 \times 10^{-4}, 0, 0)$ , and the corresponding eigenvectors

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1.00 \\ -0.260 & 0 & 4.87 \times 10^{-4} & 0.769 & 0.583 & 3.16 \times 10^{-2} & 0 \\ 0.253 & -1.13 \times 10^{-5} & -6.56 \times 10^{-4} & 0.637 & -0.728 & 1.90 \times 10^{-2} & 0 \\ 0.870 & -1.85 \times 10^{-5} & -1.07 \times 10^{-3} & 2.58 \times 10^{-2} & 0.335 & 0.361 & 0 \\ -0.334 & 2.46 \times 10^{-4} & 1.43 \times 10^{-2} & -4.91 \times 10^{-2} & -0.135 & 0.931 & 0 \\ 6.01 \times 10^{-3} & 1.75 \times 10^{-2} & 1.00 & 7.74 \times 10^{-4} & 1.53 \times 10^{-3} & -1.30 \times 10^{-2} & 0 \\ 0 & -1.00 & 1.75 \times 10^{-2} & 0 & 0 & 0 & 0 \end{pmatrix} \quad (22)$$

Values whose magnitude was less than  $10^{-5}$ , were set to zero. Two eigenvalues of  $\mathbf{Y}$  are zero. This implies that aggregation is exact for lumped systems of dimension 6. The lumping matrix of dimension  $m$  is constructed with the first  $m$  eigenvectors.

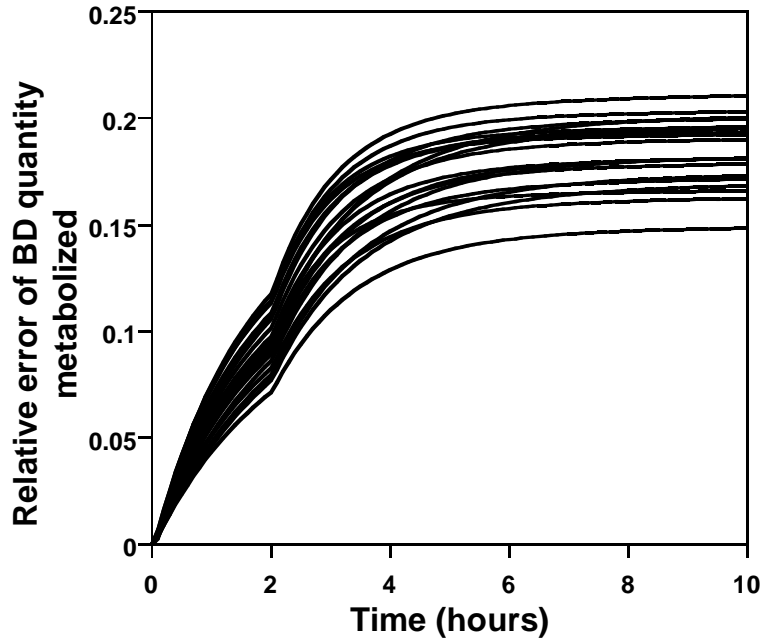
#### *Effect of lumping on the determination of the quantity of BD metabolized*

Different dimensions were tested for the lumped model. Figure 2 shows the results of constrained lumping on the BD model. The predicted BD quantity metabolized is represented for systems with dimension 6 to 2 (100 points were used to produce the graphs). Predicted quantity of BD metabolized versus time curves can be split into two parts: during exposure, the quantity metabolized increases, and then comes close to an upper bound at the end of inhalation. As expected, the lower the model dimension, the higher the error. Yet, the sign of errors can also change abruptly. Eight hours after exposure, the solution of the system with dimension 6 does not differ by more than 0.3% from the one of the initial system. For a system of dimension 5, the maximum relative error is 20%. Error goes up to 93% for dimension 3, and to 138% for dimension 2.



**Figure 2:** Time evolution of the quantity of butadiene metabolized, after an exposure to 5 ppm for 2 hours, predicted with the original system (plain curve) and lumped systems with dimension 6 to 2. 100 points were used to produce these graphs. Lumped systems were defined under the constraint that the total quantity of butadiene metabolized was left unlumped. The plain curve with the crosses corresponds to the prediction for the system of dimension 6 (which overlays the solution of the original model), the long dashed curve to the system of dimension 5, the short dashed curve to the system of dimension 4, the dotted curve to the system of dimension 3, and the dotted curve with crosses to the system of dimension 2.

The above results were obtained by using only one set of parameter values. To confirm that dimension reduction is robust to uncertainty in parameter values, different parameter sets were used and obtained by Monte Carlo simulations. These values were sampled from normal distribution of the parameters centered on the value used initially with coefficient of variation of 10%. For 20 parameter vectors, the relative error between the initial system and the lumped system was then calculated as a function of time. For the lumped system of dimension 6, the maximum error is 0.5% 8 hours after exposure ( $t = 10$ ). Figure 3 shows the time evolution of the relative error between the prediction of  $Q_{Met}$  with original system and the prediction of the lumped system of dimension 5, for 20 parameter vectors. The maximum error lies between 14% and 22%, at  $t = 10$  (i.e., the end of the simulations).



**Figure 3:** Time evolution of the relative error of the butadiene quantity metabolized, after an exposure to 5 ppm for 2 hours, between the prediction of the original system and the prediction of the lumped system of dimension 5 (the total quantity of butadiene metabolized was left unlumped). Twenty parameter vectors were used.

#### Changing the constraints

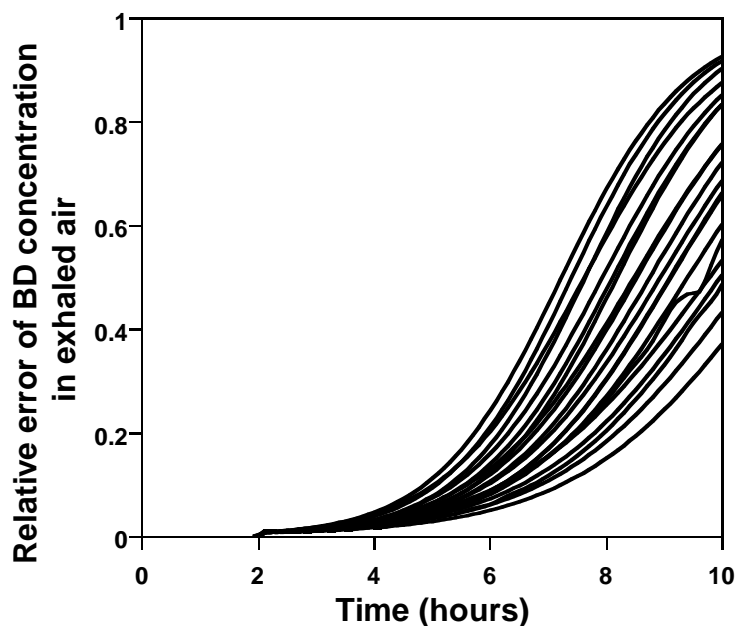
Other constraints can be applied. For example, toxicokinetic models are typically developed to be fitted to experimental data (even PBPK models, as demonstrated by Gelman *et al.*, 1996b). If needed, the lumped model may retain the possibility to be fitted to data. This implies that the model variables for which experimental data are available should not be lumped with others. In the study conducted by Bois *et al.* (1999b) (reference for our exposure scenario), exhaled air measurements were performed for each human volunteer. Given the eq. 18, we hence impose that the state variable standing for the BD quantity in lungs,  $Q_{Lu}$ , is left unlumped, in addition to the BD quantity metabolized ( $Q_{Met}$ ). The sub-matrix  $\mathbf{M}_G$  containing the constraints is therefore defined by

$$\mathbf{M}_G = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad (23)$$

The same calculations as above were performed. For the lumped system of dimension 6, the predicted quantity metabolized does not differ by more than 0.5% (at  $t = 10$ ) from the prediction of the original system. The deviations of the predicted BD concentration in exhaled air from the original solution are presented on Figure 4. The relative error increases rapidly with the observation time. Four hours after the beginning of exposure ( $t = 4$ ), this error is less than 4%. It increases to 10% one hour later ( $t = 5$ ), and it goes to 90% at

$t = 10$ . Given these results, how to conclude on the utility of the lumped system of dimension 6 ? The increase in relative error for  $C_{exh}$  is linked to the fact that the concentration of BD in exhaled air decreases rapidly. At the end of exposure ( $t = 2$ ),  $C_{exh}$  is equal to 4.6 ppm, at  $t = 4$   $C_{exh}$  is 0.023 ppm, at  $t = 5$   $C_{exh}$  is 0.010 ppm, and at  $t = 10$   $C_{exh}$  is 0.0008 ppm. At  $t = 6$ ,  $C_{exh}$  is already under the limit of quantitation, *i.e.* 0.006 ppm (Smith et al., 2001). In this example,  $C_{exh}$  is left unlumped in order to fit the lumped model. Since experimental data could not be obtained 6 hours after the beginning of exposure, performance of the lumped model after that time is irrelevant. We can therefore define a criterion which accepts the lumped model if the relative error of its predictions at the limit of quantitation does not exceed a tenth (for example) of the relative error of measurements at the same limit (which is about 200%). At  $t = 5$ , the relative error induced by the lumped system is 10%, which fulfils the criterion. We can therefore consider that the lumped model of dimension 6 can be used, and that the loss of information (compared to the original system) is minimal.

For the lumped system of dimension 5, we obtained a maximum relative error for  $Q_{Met}$  equivalent to the one of the first lumping scheme, that is around 20%, which might be deemed acceptable. However, the prediction of  $C_{exh}$  is not at all satisfying according to the above criterion. During exposure ( $t$  between 0 and 2), the error is less than 5%. But at  $t = 5$ ,  $C_{exh}$  predicted by the lumped system tends rapidly to zero, and the maximum relative error is equal to 100%, which is not acceptable according to our criterion.



**Figure 4:** Time evolution of the relative error of the concentration of butadiene in exhaled air, after an exposure to 5 ppm for 2 hours, between the prediction of the original system and the prediction of the lumped system of dimension 6 (the total quantity of butadiene metabolized and the quantity in lungs were left unlumped). Twenty parameter vectors were used.

## DISCUSSION

In this work, we have presented and applied mathematical lumping methods to toxico/pharmacokinetic models. These approaches have been initially proposed by Wei and Kuo (1969), and improved later by Li and Rabitz (1989; 1990; 1991b; 1991c; 1993; 1997). To lump compartments of a PK model, Nestorov *et al.* (1998) have proposed a method mainly based on the biochemical properties of the substance, *i.e.*, on the values taken by the model parameters (matrix **K**). The approach developed by Li and Rabitz (the one used here), is based on a formal mathematical analysis of the entire system, to lump variables that can approximate the predictions for the original or explicit variables. That approach has the potential to reveal similarities in variables that may not be apparent in an analysis solely based on considerations of their attributes (Wang *et al.*, 1998). These lumping techniques are applicable to any kind of differential equation systems, and can be easily automated. Moreover, these methods can be systematized in the exploration of optimal possible lumping. In practice, however, we need the lumped system to meet predefined (or precise) goals. To that aim, some state variables (concentrations) should not be lumped with others. With such constraints, the system cannot be usually lumped without changing to some extent the behavior of the model (and hence loosing accuracy in its predictions).



We have first treated a 2-compartment model by symbolic lumping. Reduced systems obtained by symbolic lumping are valid for all possible parameter values satisfying conditions specified in eq. 2. Numerical lumping follows the same rules but works with a fully specified original model (including numerical values of the elements of the coefficient matrix  $\mathbf{K}$ ). Calculations can then be done numerically, but the lumped systems obtained are specific to the parameter values used and hence to the chemical. Obviously, symbolic lumping is more general and flexible than numerical lumping. However the calculation of eigenvalues and eigenvectors cannot usually be performed symbolically for large dimensional matrices/models. Typically, low dimensional models (such as data-based compartmental PK models) can be treated symbolically as well as numerically, whereas high dimensional models (such as PBPK ones) are best treated numerically. However, large systems having special symmetries may be lumped symbolically. A compromise, between these two approaches, is to set some parameters and lump (if possible) the system with few symbolic parameters.

We also treated the reduction of a 6-compartment PBPK model for 1,3-butadiene with numerical constrained lumping. For this example, we have imposed two constraints on the state variables of the new lumped system: that the quantity of butadiene metabolized ( $Q_{Met}$ ), and the concentration in exhaled air ( $C_{exh}$ ) should be left unlumped. Only  $Q_{Met}$  was unlumped in scenario I, and both were unlumped in scenario II. In this case, model reduction led to a set of approximate models. Results obtained with these lumped models were tested to check if they gave appropriate answers to actual toxicokinetic questions. To insure that the lumped model is appropriate, it is necessary to define a quantitative criterion for each value of interest. In the BD example, our objective was to quantify the amount of BD metabolized after an exposure. For scenario I and II, we have shown that the original system of dimension 7 can be successfully reduced into a 6-dimensional system (less than 1% of information for  $Q_{Met}$  is lost). The reduction into a 5-dimensional system is not really satisfying for both scenarios. Possible model reduction in the above example is not large. With mathematical lumping methods, the model was reduced by only one dimension (from seven, initially). We can therefore conclude that all but one compartments in the original system are necessary to correctly describe BD biodistribution, and to quantify its fraction metabolized.

Physiological pharmacokinetic models and other *in silico* tools for fast throughput screening will certainly undergo a development and increase in complexity in the future. While adding complexity to a model is a relatively easy heuristic task, simplifying a complex model is typically more difficult, since that is in itself a complicated problem. Formal lumping techniques, such as the one presented here, have the potential to automatically reduce models and tailor them to specific needs (statistical validation versus predictions, etc.). The process could be made transparent to the user and information gained from the reduced model could be transferred back to the full model via Bayesian updating, for example.

## REFERENCES

- Bellman, R. (1960). *Introduction to Matrix Analysis*. McGraw-Hill Book Company, Inc., New York, Toronto, London.
- Bjorkman, S. (2003). Reduction and lumping of physiologically based pharmacokinetic models: Prediction of the disposition of fentanyl and pethidine in humans by successively simplified models. *J. Pharmacokinet. Pharmacodyn.* 30, 285-307.
- Bois, F. Y., Smith, T. J., Gelman, A., Chang, H. Y., and Smith, A. E. (1999). Optimal design for a study of butadiene toxicokinetics in humans. *Toxicol. Sci.* 49, 213-224.
- Daldrup, H. E., Shames, D. M., Hussein, W., Wendland, M. F., Okuhata, Y., and Brasch, R. C. (1998). Quantification of the extraction fraction for gadopentetate across breast cancer capillaries. *Magnetic Resonance in Medicine* 40, 537-543.
- Dennison, J. E., Andersen, M. E., Dobrev, I. D., Mumtaz, M. M., and Yang, R. S. H. (2004). PBPK modeling of complex hydrocarbon mixtures: gasoline. *Environmental Toxicology and Pharmacology* 16, 107-119.
- Dennison, J. E., Andersen, M. E., and Yang, R. S. H. (2003). Characterization of the pharmacokinetics of gasoline using PBPK modeling with a complex mixtures chemical lumping approach. *Inhal. Toxicol.* 15, 961-986.
- Gelman, A., Bois, F. Y., and Jiang, J. (1996). Physiological pharmacokinetic analysis using population modeling and informative prior distributions. *J. Am. Stat. Assoc.* 91, 1400-1412.
- Gibaldi, M., and Perrier, D. (1982). *Pharmacokinetics*. Marcel Dekker, New York.
- Jacquez, J. A., and Simon, C. P. (1993). Qualitative theory of compartmental systems. *Siam Review* 35, 43-79.
- Kohn, M. C., and Melnick, R. L. (1993). Species differences in the production and clearance of 1,3-butadiene metabolites: a mechanistic model indicates predominantly physiological, not biochemical, control. *Carcinogenesis* 14, 619-628.
- Li, G., and Rabitz, H. (1989). A general analysis of exact lumping in chemical kinetics. *Chemical Engineering Science* 44, 1413-1430.
- Li, G., and Rabitz, H. (1990). A general analysis of approximate lumping in chemical kinetics. *Chemical Engineering Science* 45, 977-1002.
- Li, G., and Rabitz, H. (1991a). Determination of constrained lumping schemes for nonisothermal first-order reaction systems. *Chemical Engineering Science* 46, 583-596.
- Li, G., and Rabitz, H. (1991b). A general lumping analysis of a reaction system coupled with diffusion. *Chemical Engineering Science* 46, 2041-2053.
- Li, G., and Rabitz, H. (1991c). New approaches to determination of constrained lumping schemes for a reaction system in the whole composition space. *Chemical Engineering Science* 46, 95-111.

Li, G., and Rabitz, H. (1993). The direct lumping approach: an application to a catalytic reforming model. *Chemical Engineering Science* 48, 1903-1909.

Li, G., and Rabitz, H. (1997). Reduced kinetic equations of a CO/H<sub>2</sub>/air oxidation model by a special perturbation method. *Chemical Engineering Science* 52, 4317-4327.

Li, G., Rabitz, H., and Toth, J. (1994a). A general analysis of exact nonlinear lumping in chemical kinetics. *Chemical Engineering Science* 49, 343-361.

Li, G. Y., Tomlin, A. S., Rabitz, H., and Toth, J. (1994b). A general analysis of approximate nonlinear lumping in chemical kinetics. II. Unconstrained lumping. *Journal of Chemical Physics* 101, 1172-1187.

Nestorov, I. A., Aarons, L. J., Arundel, P. A., and Rowland, M. (1998). Lumping of whole-body physiologically based pharmacokinetic models. *J. Pharmacokinet. Biopharm.* 26, 21-46.

Smith, T., Lin, Y.-S., Mezzetti, L., Bois, F. Y., Kelsey, K., and Ibrahim, J. (2001). Genetic and dietary factors affecting human metabolism of 1,3-butadiene. *Chemico-Biological Interactions* 135-136, 407-428.

Tomlin, A. S., Turanyi, T., and Pilling, M. J. (1997). Mathematical tools for the construction, investigation and reduction of combustion mechanisms. In *Low-temperature combustion and autoignition* (M. J. Pilling, ed., Vol. 35, pp. 293-437. Elsevier, Amsterdam.

Verhaar, H. J., Morroni, J. R., Reardon, K. F., Hays, S. M., Gaver, D. P., Jr., Carpenter, R. L., and Yang, R. S. (1997). A proposed approach to study the toxicology of complex mixtures of petroleum products: the integrated use of QSAR, lumping analysis and PBPK/PD modeling. *Environmental Health Perspectives* 105 Suppl 1, 179-95.

Wang, S. W., Georgopoulos, P. G., Li, G., and Rabitz, H. (1998). Condensing complex atmospheric chemistry mechanisms. 1. The direct constrained approximate lumping (DCAL) method applied to alkane photochemistry. *Environmental Science & Technology* 32, 2018-2024.

Wei, J., and Kuo, J. C. W. (1969). A lumping analysis in monomolecular reaction systems. Analysis of the exactly lumpable system. *Industrial & Engineering chemistry fundamentals* 8, 114-123.

## 7. Conclusion et perspectives

Dans ce travail, nous avons développé deux protocoles permettant de paramétrer les modèles physiologiques toxico/pharmacocinétiques : le couplage de la modélisation PBPK avec des données d'imagerie, et l'utilisation de traceurs "explorateurs" dans les études toxicocinétiques. Les deux protocoles que nous proposons ont pour objectif d'intégrer de nouveaux types de données (telles que les concentrations internes), et de renforcer la connaissance sur la physiologie du sujet. Pour chacune des analyses effectuées, nous nous sommes placés dans un cadre statistique bayésien. Ce cadre inférentiel est particulièrement adapté pour l'analyse des modèles PBPK, car il permet d'intégrer de l'information connue *a priori* dans le processus d'estimation des paramètres et d'estimer l'incertitude sur les estimations.

Dans une première partie, nous avons présenté l'analyse statistique bayésienne de données toxicocinétiques. L'objectif de ce travail était de définir les différentes étapes nécessaires à la construction d'un modèle toxico/pharmacocinétique, et de montrer l'utilité de la statistique bayésienne pour calibrer ces modèles. Une des conclusions de cette partie est qu'il est nécessaire de modéliser rigoureusement les incertitudes inhérentes aux données toxicocinétiques (erreur de mesure, variabilité dans la population, etc.) pour décrire au mieux la cinétique de la substance d'intérêt. Nous montrons que le cadre inférentiel bayésien est particulièrement adapté pour l'analyse de données toxicocinétiques.

Dans la seconde partie, nous avons appliqué l'analyse d'image au développement d'un modèle PBPK décrivant la biodistribution d'un agent de contraste, Vistarem<sup>®</sup>. Grâce à la modélisation des données recueillies par des techniques d'IRM, nous avons pu caractériser la vascularisation du corps entier (en particulier, la perfusion, la perméabilité, et la fraction volumique des capillaires), et celle d'une tumeur greffée. Il fut ainsi possible d'évaluer statistiquement les effets d'un traitement antiangiogénique (c'est-à-dire inhibant la formation de nouveaux vaisseaux), en comparant un groupe d'animaux traités et un groupe placebo. Au-delà de ces résultats, nous avons montré qu'il est possible d'identifier un phénomène physiologique (le réseau vasculaire) en couplant des données recueillies par IRM et un modèle physiologique. Une détermination précise du réseau vasculaire est de première importance pour la modélisation PBPK, puisque le sang est le principal véhicule de nombreuses substances dans l'organisme. La prédiction de la pharmacocinétique d'une substance toxique ou thérapeutique serait donc potentiellement améliorée avec une caractérisation précise du réseau vasculaire.

Dans la troisième partie, nous avons introduit l'utilisation de traceurs explorateurs lors d'études toxicocinétiques. Nous avons développé des protocoles d'exposition à une substance chimique et à un traceur non-toxique pour les humains. Notre objectif était de paramétrer partiellement le modèle PBPK à l'aide des données expérimentales sur le traceur, puis de compléter cette paramétrisation à l'aide d'un minimum de données sur la substance toxique d'intérêt. Sur la base d'expositions simulées, nous avons montré qu'en administrant aux sujets deux substances, il est possible de réduire la dose d'exposition du toxique (par 4 dans notre exemple d'application) tout en maintenant la précision sur l'estimation des paramètres d'intérêt.

Malgré un nombre réduit de données disponibles sur la substance toxique, notre protocole devrait donc permettre de caractériser au mieux sa toxicocinétique.

Un des inconvénients de ce protocole d'exposition est la lourdeur des calculs induits. En effet, la calibration de deux modèles PBPK sont nécessaires, un pour la substance toxique et un pour le traceur. Afin de réduire la complexité des calculs engendrés, nous avons proposé d'appliquer des techniques d'agrégation de systèmes d'équations différentielles. L'essentiel de ces techniques est de réduire un système tout en conservant sa dynamique. Nous avons choisi d'appliquer la méthode développée par Li et Rabitz (1989), qui permet d'agréger un système d'équations différentielles en se basant sur sa structure mathématique. L'avantage principal de cette méthode est d'être généraliste (c'est-à-dire de ne pas dépendre du système lui-même). Des améliorations lui ont été apportées pour permettre aux utilisateurs d'imposer des contraintes sur le nouveau système (Li and Rabitz, 1991c). Ces techniques ont été initialement développées pour l'industrie pétrolière, nous avons proposé de les appliquer à des modèles toxico/pharmacocinétiques. Nous avons pu réduire des modèles relativement simples, afin de montrer l'utilité des méthodes d'agrégation dans ce domaine. Ces techniques nécessitent cependant encore des recherches. La principale amélioration à apporter est la possibilité d'appliquer ces techniques sur des systèmes analytiques, et non numériques. Pour des exemples simples, des solutions analytiques peuvent être obtenues. Mais, en pratique, ces exemples restent relativement rares.

Chacun des deux protocoles développés dans ce travail inclut, en plus de la substance d'intérêt, l'utilisation d'une substance "annexe" (un agent de contraste ou un traceur explorateur). Notre travail montre que l'utilisation de cette substance annexe (ou traceur) permet de renforcer la connaissance sur l'anatomie et la physiologie du sujet considéré. Ceci conduit donc à une meilleure détermination de l'action (toxique ou bénéfique) de la substance d'intérêt. Le premier protocole étudie le traceur par des techniques d'imagerie, alors que le second l'étudie par l'intermédiaire de la pharmacocinétique classique. Il serait cependant possible de coupler ces deux protocoles. Une suite logique de notre travail serait de développer des protocoles d'exposition à une substance toxique (telle que le 1,3-butadiène) et à une substance visualisable par une technique d'imagerie quantitative (telle qu'un agent de contraste). Un tel protocole requiert cependant un développement relativement long, et des moyens techniques importants. Mais il pourrait se révéler utile pour des substances pour lesquelles peu de données sur le toxique sont disponibles.

Dans la suite, nous présentons divers axes de recherche qui pourraient bénéficier de nos protocoles.

#### *Réduire les risques des expérimentations chez l'humain*

Les études animales sont typiquement utilisées pour prédire le comportement de substances chimiques chez l'humain. Cependant, pour certaines substances, ces études sont inappropriées pour cet usage, lorsque les effets diffèrent entre espèces (Olson *et al.*, 2000). Dans ces cas précis, l'US EPA juge qu'il est prioritaire de conduire des études humaines, afin de pouvoir prendre des décisions pertinentes pour la protection des populations (Dourson *et al.*, 2001; US Environmental Protection Agency, 2000). Pour des considérations éthiques évidentes, il est toutefois essentiel que ces expositions n'engendrent qu'un risque négligeable pour

les volontaires humains. Dans ce contexte, il est donc nécessaire de développer des techniques réduisant au maximum les doses d'exposition lors de ces études, tout en permettant la caractérisation des risques liés à l'exposition. Les protocoles développés ici s'inscrivent directement dans le cadre de ces techniques. Nous montrons qu'en exposant les sujets à une substance traceur n'ayant pas de conséquence néfaste, il est possible de mieux caractériser la cinétique d'une substance (en terme de précision des estimations des paramètres du modèle, par exemple).

#### *Valider les modèles PKPB et les données*

En plus de données pertinentes, il est nécessaire de développer des modèles adéquats. Les données fournies par les techniques d'imagerie quantitative pourraient renforcer la fiabilité des modèles PBPK en permettant, notamment, de les valider. On distingue deux types de validation (Kohn, 1995) : la validation des données anatomiques, physiologiques et biochimiques, et la validation de la structure mathématique du modèle.

Dans notre travail, nous avons montré qu'il est possible, avec des données d'IRM, d'obtenir des informations sur les caractéristiques pharmacocinétiques du sujet de manière non-invasive. Dans notre cas, des informations individuelles précises sur le réseau vasculaire d'une tumeur ont été obtenues. En diversifiant les mesures de concentrations de l'agent de contraste, il aurait été possible d'obtenir ces mêmes informations pour divers tissus ou organes. Le développement de ce type de données peut facilement s'intégrer dans les différents travaux visant à créer des bases de données regroupant les valeurs de paramètres nécessaires à la modélisation PBPK et leur distribution aléatoire dans la population pour différentes espèces (Price *et al.*, 2003). La création de telles bases aurait pour objectif d'éviter la redondance d'expérimentations (animales, par exemple), et de permettre l'uniformisation des valeurs de paramètres utilisées.

Pour valider la structure mathématique d'un modèle PBPK, de nombreux auteurs disent qu'il suffit uniquement de s'assurer que les prédictions du modèle concordent avec les données expérimentales observées (Gargas *et al.*, 2000b; Gargas *et al.*, 2000d; Haddad *et al.*, 2000; Mann *et al.*, 1996). Cependant, les seules données toxicocinétiques disponibles sont typiquement les concentrations dans le sang, les urines ou l'air exhalé. De même, il est rare que plusieurs types de données soient développés pour une même étude. Kohn (1997) a testé la pertinence de ce type de validation en comparant les prédictions de trois différents modèles PBPK à un même ensemble de données. Les trois modèles reproduisent correctement les données disponibles, mais fournissent des prédictions différentes pour les concentrations dans les autres compartiments. Il n'est donc pas possible que tous ces modèles soient corrects, et les données habituelles ne suffisent pas à les discriminer. La validation de ces modèles requiert donc bien plus (Kohn, 1995, 1997). Afin de vérifier la fiabilité des prédictions des doses tissulaires internes d'un modèle, il est donc essentiel de continuer le développement de méthodes expérimentales permettant la détermination de ces doses *in vivo* ou dans les cellules *in vitro* (Andersen, 2003). Comme nous le proposons dans ce travail, un des axes de progrès serait le développement des techniques d'imagerie permettant la quantification des concentrations internes.

*Réduire la dimension des modèles*

Les techniques d'agrégation ont un fort potentiel pour la modélisation physiologique. Un exemple intéressant est la modélisation des mélanges de substances chimiques (Andersen and Dennison, 2004; Dennison *et al.*, 2004; Simmons, 1996). Dans la vie courante, nous ne sommes pas exposés à une substance à la fois, mais à des mélanges. Pour quantifier les risques encourus lors d'exposition à des mélanges, un modèle est défini pour chaque substance, et ensuite leurs interactions sont ensuite modélisées. La dimension du système à analyser peut donc rapidement croître. Cependant la manipulation mathématique de ces modèles de grande dimension est souvent problématique, et pourrait bénéficier des techniques d'agrégation.

*Organiser les connaissances*

En plus d'être des outils de prédiction d'effets toxiques ou bénéfiques, les modèles PBPK peuvent s'intégrer à des projets plus globaux. Actuellement, plusieurs groupes travaillent sur le développement de projets rassemblant diverses sources de connaissance sur les êtres humains (leur anatomie, physiologie, génétique, etc.) afin de fournir des outils permettant de simuler virtuellement un être humain, ou d'y prédire le comportement de substances dans cet organisme (Timchalk *et al.*, 2001). On peut citer, par exemple, le "Virtual Human Program" (Ward *et al.*, 2000) ou le "Physiome Project" (Bassingthwaighe, 2000a, b; Kohl *et al.*, 2000). Ces projets ambitieux vont donc nécessiter le développement de modèles biologiquement fondés permettant de rassembler, organiser et expliquer les diverses sources de données. Il est fondamental de travailler à l'analyse quantitative de ces données.

## 8. Abréviations

ADME	Administration, Distribution, Métabolisme, Excrétion
BUGS	Bayesian Inference Using Gibbs Sampler
DCAL	Direct Constrained Approximate Lumping
IRM	Imagerie par Résonance Magnétique
MCMC	Monte Carlo Markov Chain
MCSim	Monte Carlo Simulation
NOAEL	No-Observed-Adverse-Effect Level
QSAR	Quantitative Structure-Activity Relationship
RMN	Résonance Magnétique Nucléaire
PBPK	Physiologically Based Pharmacokinetic
TEP	Tomographie par Emission de Positrons



## 9. Références

- Aarons, L. (1992). Population pharmacokinetics. *International Journal of Clinical Pharmacology, Therapy, and Toxicology* **30**, 520-522.
- Aine, C. J. (1995). A conceptual overview and critique of functional neuroimaging techniques in humans: I. MRI/fMRI and PET. *Critical Reviews in Neurobiology* **9**, 229-309.
- Amzal, B., Bois, F. Y., Parent, E., and Robert, C. P. (2004). Bayesian Optimal Design via Interacting MCMC. *Cahiers du Ceremade* **034**.
- Andersen, M. E. (1991). Physiological modelling of organic compounds. *The Annals of Occupational Hygiene* **35**, 309-321.
- Andersen, M. E. (1995a). Development of physiologically based pharmacokinetic and physiologically based pharmacodynamic models for applications in toxicology and risk assessment. *Toxicology Letters* **79**, 35-44.
- Andersen, M. E. (1995b). Physiologically based pharmacokinetic (PB-PK) models in the study of the disposition and biological effects of xenobiotics and drugs. *Toxicology Letters* **82-83**, 341-348.
- Andersen, M. E. (2003). Toxicokinetic modeling and its applications in chemical risk assessment. *Toxicology Letters* **138**, 9-27.
- Andersen, M. E., and Dennison, J. E. (2004). Mechanistic approaches for mixture risk assessments - Present capabilities with simple mixtures and future directions. *Environmental Toxicology and Pharmacology* **16**, 1-11.
- Ando, Y., Fukatsu, H., Ishiguchi, T., Ishigaki, T., Endo, T., and Miyazaki, M. (2000). Diagnostic utility of tumor vascularity on magnetic resonance imaging of the breast. *Magnetic Resonance Imaging* **18**, 807-813.
- Ardekani, B. A., and Kanno, I. (1998). Statistical methods for detecting activated regions in functional MRI of the brain. *Magnetic Resonance Imaging* **16**, 1217-1225.
- Arms, A. D., and Travis, C. C. (1988). Reference Physiological Parameters in Pharmacokinetics modeling. U.S. Environmental Protection Agency.
- Bailer, A. J., and Dankovic, D. A. (1997). An introduction to the use of physiologically based pharmacokinetic models in risk assessment. *Statistical Methods in Medical Research* **6**, 341-358.
- Balant, L. P., and Gex-Fabry, M. (1990). Physiological pharmacokinetic modelling. *Xenobiotica* **20**, 1241-1257.
- Balant, L. P., and Gex-Fabry, M. (2000). Modelling during drug development. *European Journal of Pharmaceutics and Biopharmaceutics* **50**, 13-26.
- Bassingthwaight, J. B. (2000a). Back to fundamentals: Anatomy-based physiological bioengineering. *Annals of Biomedical Engineering* **28**, 701-703.

- Bassingthwaight, J. B. (2000b). Strategies for the physiome project. *Annals of Biomedical Engineering* **28**, 1043-1058.
- Bayard, D. S., and Jelliffe, R. W. (2004). A Bayesian approach to tracking patients having changing pharmacokinetic parameters. *Journal of Pharmacokinetics and Pharmacodynamics* **31**, 75-107.
- Beal, S. L., and Sheiner, L. B. (1982). Estimating population kinetics. *CRC Critical Reviews in Biomedical Engineering* **8**, 195-222.
- Beck, B. D., and Clewell, H. J. (2001). Uncertainty/Safety Factors in Health Risk Assessment: Opportunities for Improvement. *Human and Ecological Risk Assessment* **7**, 203-207.
- Beckmann, N., Mueggler, T., Allegrini, P. R., Laurent, D., and Rudin, M. (2001). From anatomy to the target: Contributions of magnetic resonance imaging to preclinical pharmaceutical research. *Anatomical Record* **265**, 85-100.
- Bellman, R. (1960). *Introduction to Matrix Analysis*. McGraw-Hill Book Company, Inc., New York, Toronto, London.
- Bernardo, J. M., and Smith, A. F. M. (1994). *Bayesian Theory*. Wiley, New York.
- Bernillon, P. (2000). Développement de modèles toxicocinétiques et analyse statistique de données d'exposition au tétrachloroéthylène. Thèse de l'université Pierre et Marie Curie, Paris.
- Bernillon, P., and Bois, F. Y. (2000). Statistical issues in toxicokinetic modeling: A Bayesian perspective. *Environmental Health Perspectives* **108**, 883-893.
- Berns, G. S. (1999). Functional Neuroimaging. *Life Sciences* **65**, 2531-2540.
- Bertrand, M., Jackson, P., and Walther, B. (2000). Rapid assessment of drug metabolism in the drug discovery process. *European Journal of Pharmaceutical Sciences* **11**, S61-S72.
- Bischoff, K. B., and Brown, R. G. (1966). drug distribution in mammals. *Chemical Engineering Progress* **66**, 33-45.
- Bischoff, K. B., Dedrick, R. L., Zaharko, D. S., and Longstreth, J. A. (1971). Methotrexate pharmacokinetics. *Journal of Pharmaceutical Sciences* **60**, 1128-1133.
- Bjorkman, S. (2003). Reduction and lumping of physiologically based pharmacokinetic models: Prediction of the disposition of fentanyl and pethidine in humans by successively simplified models. *Journal of Pharmacokinetics and Pharmacodynamics* **30**, 285-307.
- Blesch, K. S., Gieschke, R., Tsukamoto, Y., Reigner, B. G., Burger, H. U., and Steimer, J. L. (2003). Clinical pharmacokinetic/pharmacodynamic and physiologically based pharmacokinetic modeling in new drug development: The capecitabine experience. *Investigational New Drugs* **21**, 195-223.
- Bois, F. Y. (1999). Analysis of PBPK models for risk characterization. *Annals of the New York Academy of Sciences* **895**, 317-337.
- Bois, F. Y. (2000). Statistical analysis of Fisher et al. PBPK model of trichloroethylene kinetics. *Environmental Health Perspectives* **108** (suppl. 2), 275-282.
- Bois, F. Y. (2001). Applications of population approaches in toxicology. *Toxicology Letters* **120**, 385-394.

- Bois, F. Y. (2003). Modélisation toxicocinétique de la concentration sanguine de 2,3,7,8-tetrachloro-p-dioxine après ingestion chez la femme. *Environnement, Risques et Santé* **2**.
- Bois, F. Y., Gelman, A., Jiang, J., Maszle, D., Zeise, L., and Alexeef, G. (1996a). Population toxicokinetics of tetrachloroethylene. *Archives of Toxicology* **70**, 347-355.
- Bois, F. Y., Gelman, A., Jiang, J., Maszle, D. R., Zeise, L., and Alexeef, G. (1996b). Population toxicokinetics of tetrachloroethylene. *Archives of Toxicology* **70**, 347-355.
- Bois, F. Y., Jackson, E., Pekari, K., and Smith, M. (1996c). Population toxicokinetics of benzene. *Environmental Health Perspectives* **104** (suppl. 6), 1405-1411.
- Bois, F. Y., and Maszle, D. (1997a). MCSim: a simulation program. *Journal of Statistical Software* **2**(9), [http://toxi.ineris.fr/activites/toxicologie\\_quantitative/mcsim/mcsim.php](http://toxi.ineris.fr/activites/toxicologie_quantitative/mcsim/mcsim.php).
- Bois, F. Y., and Maszle, D. (1997b). MCSim: a simulation program. *Journal of Statistical Software* **2**(9), <http://www.stat.ucla.edu/journals/jss/v02/i09> (software available online at [www.toxi.ineris.fr](http://www.toxi.ineris.fr)).
- Bois, F. Y., and Maszle, D. R. (1997c). MCSim: A Monte Carlo Simulation Program. available at [http://toxi.ineris.fr/activites/toxicologie\\_quantitative/mcsim/mcsim.php](http://toxi.ineris.fr/activites/toxicologie_quantitative/mcsim/mcsim.php).
- Bois, F. Y., Smith, T. J., Gelman, A., Chang, H. Y., and Smith, A. E. (1999a). Optimal design for a study of butadiene toxicokinetics in humans. *Toxicological Sciences: an Official Journal of the Society of Toxicology* **49**, 213-224.
- Bois, F. Y., Smith, T. J., Gelman, A., Chang, H. Y., and Smith, A. E. (1999b). Optimal design for a study of butadiene toxicokinetics in humans. *Toxicological Sciences* **49**, 213-224.
- Bois, F. Y., Tozer, T. N., Hauck, W. W., Chen, M. L., Patnaik, R., and Williams, R. L. (1994). Bioequivalence: performance of several measures of extent of absorption. *Pharmaceutical Research* **11**, 715-22.
- Bortot, P., Thomaseth, K., and Salvan, A. (2002). Population toxicokinetic analysis of 2,3,7,8-tetrachlorodibenzo-p-dioxin using Bayesian techniques. *Statistics in Medicine* **21**, 533-547.
- Brasch, R. C., Daldrup, H., Shames, D., Wendland, M., Okuhata, Y., and Rosenau, W. (1998). Macromolecular contrast media-enhanced MRI estimates of microvascular permeability correlate with histopathologic tumor grade. *Academic Radiology* **5** Supplement 1, S2-S5.
- Brix, G., Semmler, W., Port, R., Schad, L. R., Layer, G., and Lorenz, W. J. (1991). Pharmacokinetic parameters in CNS Gd-DTPA enhanced MR imaging. *Journal of Computer Assisted Tomography* **15**, 621-8.
- Brochot, C., Marchand, S., Couet, W., Gelman, A., and Bois, F. Y. (2004). Extension of the isobolographic approach to interactions studies between more than two drugs: Illustration with the convulsant interaction between pefloxacin, norfloxacin, and theophylline in rats. *Journal of Pharmaceutical Sciences* **93**, 553-562.
- Brown, R. P., Delp, M. D., Lindstedt, S. L., Rhomberg, L. R., and Beliles, R. P. (1997). Physiological parameter values for physiologically based pharmacokinetic models. *Toxicology and Industrial Health* **13**, 407-484.

Bruckner, J. V., Keys, D. A., and Fisher, J. W. (2004). The Acute Exposure Guideline Level (AEGL) program: Applications of physiologically based pharmacokinetic modeling. *Journal of Toxicology and Environmental Health-Part a-Current Issues* **67**, 621-634.

Carpenter, R. L., Eger, E. I., 2nd, Johnson, B. H., Unadkat, J. D., and Sheiner, L. B. (1986). The extent of metabolism of inhaled anesthetics in humans. *Anesthesiology* **65**, 201-5.

Charnick, S. B., Kawai, R., Nedelman, J. R., Lemaire, M., Niederberger, W., and Sato, H. (1995). Physiologically based pharmacokinetic modeling as a tool for drug development. *Journal of Pharmacokinetics and Biopharmaceutics* **23**, 217-229.

Chen, C. (2000). Validation of a population pharmacokinetic model for adjunctive lamotrigine therapy in children. *British Journal of Clinical Pharmacology* **50**, 135-145.

Claret, L., and Iliadis, A. (1996). Nonparametric density estimation applied to population pharmacokinetics. *Math Biosci* **133**, 51-68.

Clewell, H. J. (1995). The application of physiologically based pharmacokinetic modeling in human health risk assessment of hazardous substances. *Toxicology Letters* **79**, 207-217.

Corley, R. A., Markham, D. A., Banks, C., Delorme, P., Masterman, A., and Houle, J. M. (1997). Physiologically based pharmacokinetics and the dermal absorption of 2-butoxyethanol vapor by humans. *Fundamental and Applied Toxicology* **39**, 120-130.

Cowles, M. K., and Carlin, B. P. (1996). Markov chain Monte Carlo convergence diagnostics: a comparative review. *Journal of the American Statistical Association* **91**, 883-904.

Daldrup, H. E., Shames, D. M., Hussein, W., Wendland, M. F., Okuhata, Y., and Brasch, R. C. (1998). Quantification of the extraction fraction for gadopentetate across breast cancer capillaries. *Magnetic Resonance in Medicine* **40**, 537-543.

Davis, B., and Morris, T. (1993). Physiological parameters in laboratory animals and human. *Pharmaceutical Research* **10**, 1093-1095.

Dedrick, R. L., Forrester, D. D., Cannon, J. N., El Dareer, S. M., and Mellett, L. B. (1973). Pharmacokinetics of 1-beta-D-arabinofuranosylcytosine (ARA-C) deamination in several species. *Biochemical Pharmacology* **22**, 2405-2417.

Dennison, J. E., Andersen, M. E., Dobrev, I. D., Mumtaz, M. M., and Yang, R. S. H. (2004). PBPK modeling of complex hydrocarbon mixtures: gasoline. *Environmental Toxicology and Pharmacology* **16**, 107-119.

Dennison, J. E., Andersen, M. E., and Yang, R. S. H. (2003). Characterization of the pharmacokinetics of gasoline using PBPK modeling with a complex mixtures chemical lumping approach. *Inhalation Toxicology* **15**, 961-986.

Deurenberg, P., Weststrate, J. A., and Seidell, J. C. (1991). Body mass index as a measure of body fatness: age- and sex-specific prediction formulas. *British Journal of Nutrition* **65**, 105-14.

Dourson, M. L., Andersen, M. E., Erdreich, L. S., and MacGregor, J. A. (2001). Using human data to protect the public's health. *Regulatory Toxicology and Pharmacology* **33**, 234-256.

Droesbeke, J., Fine, J., and Saporta, G. (2002). *Méthodes Bayésiennes en Statistiques*. Société Française de Statistique, Paris.

Evelhoch, J. L. (1999). Key factors in the acquisition of contrast kinetic data for oncology. *Journal of Magnetic Resonance Imaging* **10**, 254-9.

Evelhoch, J. L., McDouall, J. B., Mattiello, J., and Simpson, N. E. (1992). Measurement of relative regional tumor blood flow in mice by deuterium NMR imaging. *Magnetic Resonance in Medicine* **24**, 42-52.

Everitt, B. S., and Bullmore, E. T. (1999). Mixture model mapping of brain activation in functional magnetic resonance images. *Human Brain Mapping* **7**, 1-14.

Fernandez, J. G., Droz, P. O., Humbert, B. E., and Caperos, J. R. (1977). Trichloroethylene exposure - Simulation of uptake, excretion, and metabolism using a mathematical model. *British Journal of Industrial Medicine* **34**, 43-55.

Gargas, M. L., Tyler, T. R., Sweeney, L. M., Corley, R. A., Weitz, K. K., Mast, T. J., Paustenbach, D. J., and Hays, S. M. (2000a). A toxicokinetic study of inhaled ethylene glycol ethyl ether acetate and validation of a physiologically based pharmacokinetic model for rat and human. *Toxicol Appl Pharmacol* **165**, 63-73.

Gargas, M. L., Tyler, T. R., Sweeney, L. M., Corley, R. A., Weitz, K. K., Mast, T. J., Paustenbach, D. J., and Hays, S. M. (2000b). A toxicokinetic study of inhaled ethylene glycol ethyl ether acetate and validation of a physiologically based pharmacokinetic model for rat and human. *Toxicology and Applied Pharmacology* **165**, 63-73.

Gargas, M. L., Tyler, T. R., Sweeney, L. M., Corley, R. A., Weitz, K. K., Mast, T. J., Paustenbach, D. J., and Hays, S. M. (2000c). A toxicokinetic study of inhaled ethylene glycol monomethyl ether (2-ME) and validation of a physiologically based pharmacokinetic model for the pregnant rat and human. *Toxicol Appl Pharmacol* **165**, 53-62.

Gargas, M. L., Tyler, T. R., Sweeney, L. M., Corley, R. A., Weitz, K. K., Mast, T. J., Paustenbach, D. J., and Hays, S. M. (2000d). A toxicokinetic study of inhaled ethylene glycol monomethyl ether (2-ME) and validation of a physiologically based pharmacokinetic model for the pregnant rat and human. *Toxicology and Applied Pharmacology* **165**, 53-62.

Gelman, A., Bois, F. Y., and Jiang, J. (1996a). Physiological pharmacokinetic analysis using population modeling and informative prior distributions. *Journal of the American Statistical Association* **91**, 1400-1412.

Gelman, A., Bois, F. Y., and Jiang, J. (1996b). Physiological pharmacokinetic analysis using population modeling and informative prior distributions. *Journal of the American Statistical Association* **91**, 1400-1412.

Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (1995). *Bayesian Data Analysis*. Chapman & Hall, London.

Gelman, A., and Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences (with discussion). *Statistical Science* **7**, 457-511.

- Gelman, A., and Rubin, D. B. (1996). Markov chain Monte Carlo methods in biostatistics. *Statistical Methods in Medical Research* **5**, 339-355.
- Geman, S., and Geman, D. (1984). Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. *Ieee Transactions on Pattern Analysis and Machine Intelligence* **6**, 721-741.
- Genovese, C. R. (2000). A Bayesian time-course model for functional magnetic resonance imaging data. *Journal of the American Statistical Association* **95**, 691-703.
- Gerlowski, L. E., and Jain, R. K. (1983). Physiologically based pharmacokinetic modeling: principles and applications. *Journal of Pharmaceutical Sciences* **72**, 1103-1127.
- Ghafourian, T., Barzegar-Jalali, M., Hakimiha, N., and Cronin, M. T. D. (2004). Quantitative structure-pharmacokinetic relationship modelling: apparent volume of distribution. *Journal of Pharmacy and Pharmacology* **56**, 339-350.
- Gibaldi, M., and Perrier, D. (1982). *Pharmacokinetics*. Marcel Dekker, New York.
- Gilks, W. R., Richardson, S., and Spiegelhalter, D. J. (1996). *Markov Chain Monte Carlo in Practice*. Chapman & Hall, London.
- Gillies, R. J., Bhujwalla, Z. M., Evelhoch, J., Garwood, M., Neeman, M., Robinson, S. P., Sotak, C. H., and Van Der Sanden, B. (2000). Applications of magnetic resonance in model systems: tumor biology and physiology. *Neoplasia* **2**, 139-151.
- Gossl, C., Fahrmeir, L., and Auer, D. P. (2001). Bayesian modeling of the hemodynamic response function in BOLD fMRI. *Neuroimage* **14**, 140-148.
- Graves, D. A., and Chang, I. (1988). Application of NONMEM to routine bioavailability data. *Journal of Pharmacokinetics and Biopharmaceutics* **18**, 145-160.
- Haddad, S., Charest-Tardif, G., Tardif, R., and Krishnan, K. (2000). Validation of a physiological modeling framework for simulating the toxicokinetics of chemicals in mixtures. *Toxicology and Applied Pharmacology* **167**, 199-209.
- Haerberle, H. A., Wahl, H. G., Jakubetz, H., Krause, H., Schmidt, R., Schurig, V., and Dieterich, H. J. (2002). Accumulation of S(+) enantiomer in human beings after general anaesthesia with isoflurane racemate. *Eur J Anaesthesiol* **19**, 641-6.
- Hansch, C., Leo, A., Mekapati, S. B., and Kurup, A. (2004). QSAR and ADME. *Bioorganic & Medicinal Chemistry* **12**, 3391-3400.
- Hastings, W. K. (1970). Monte Carlo methods using Markov Chains and their applications. *Biometrika* **57**, 97-109.
- Hayton, P., Brady, M., Tarassenko, L., and Moore, N. (1997). Analysis of dynamic MR breast images using a model of contrast enhancement. *Medical Image Analysis* **1**, 207-24.
- Himmelstein, M. W., Turner, M. J., Asgharian, B., and Bond, J. A. (1996). Metabolism of 1,3-butadiene: inhalation pharmacokinetics and tissue dosimetry of butadiene epoxides in rats and mice. *Toxicology* **113**, 306-9.

- Ings, R. M. (1990). Interspecies scaling and comparisons in drug development and toxicokinetics. *Xenobiotica* **20**, 1201-1231.
- International Agency for Research on Cancer (1992). *Occupational Exposures to Mists and Vapours from Strong Inorganic Acids and Other Industrial Chemicals*. International Agency for Research on Cancer, Lyon, France.
- International Agency for Research on Cancer (IARC) (1992). *Occupational Exposures to Mists and Vapours from Strong Inorganic Acids and Other Industrial Chemicals*. International Agency for Research on Cancer, Lyon, France.
- International Commission on Radiological Protection (ICRP) (2002). *Basic Anatomical and Physiological data for Use in Radiological Protection: Reference Values*.
- International Commission on Radiological Protection (ICRP) (2003). *Basic Anatomical and Physiological Data for Use in Radiological Protection: Reference Values*. J. Valentin, Stockholm.
- Jacquez, J. A. (1999). *Modeling with Compartments*. BioMedware, Ann Arbor.
- Jacquez, J. A., and Simon, C. P. (1993). Qualitative theory of compartmental systems. *Siam Review* **35**, 43-79.
- Jonsson, F. (2001). Physiologically Based Pharmacokinetic modeling in risk assessment. In Division of Pharmacokinetics and Drug Therapy. Uppsala University.
- Jonsson, F., and Johanson, G. (2001). Bayesian estimation of variability in adipose tissue blood flow in man by physiologically based pharmacokinetic modeling of inhalation exposure to toluene. *Toxicology* **157**, 177-193.
- Jonsson, F., and Johanson, G. (2002). Physiologically based modeling of the inhalation kinetics of styrene in humans using a Bayesian population approach. *Toxicology and Applied Pharmacology* **179**, 35-49.
- Jonsson, F., and Johanson, G. (2003). The Bayesian population approach to physiological toxicokinetic-toxicodynamic models - an example using the MCSim software. *Toxicology Letters* **138**, 143-150.
- Kaniwa, N., Aoyagi, N., Ogata, H., and Ishii, M. (1990). Application of the NONMEM method to evaluation of the bioavailability of drug products. *Journal of Pharmaceutical Sciences* **79**, 1116-1120.
- Kashuba, A. D., Ballow, C. H., and Forrest, A. (1996). Development and evaluation of a Bayesian pharmacokinetic estimator and optimal, sparse sampling strategies for ceftazidime. *Antimicrobial Agents and Chemotherapy* **40**, 1860-5.
- Kastler, B., Vetter, D., Patay, Z., and Germain, P. (2003). *Comprendre l'IRM*. Editions Masson.
- Kohl, P., Noble, D., Winslow, R. L., and Hunter, P. J. (2000). Computational modelling of biological systems: tools and visions. *Philosophical Transactions of the Royal Society of London Series a-Mathematical Physical and Engineering Sciences* **358**, 579-610.
- Kohn, M. C. (1995). Achieving credibility in risk assessment models. *Toxicology Letters* **79**, 107-14.
- Kohn, M. C. (1997). The importance of anatomical realism for validation of physiological models of disposition of inhaled toxicants. *Toxicology and Applied Pharmacology* **147**, 448-458.

- Kohn, M. C., and Melnick, R. L. (1993). Species differences in the production and clearance of 1,3-butadiene metabolites: a mechanistic model indicates predominantly physiological, not biochemical, control. *Carcinogenesis* **14**, 619-628.
- Kohn, M. C., and Melnick, R. L. (2001). Physiological modeling of butadiene disposition in mice and rats. *Chemico-Biological Interactions* **135-136**, 285-301.
- Kuo, J. C. W., and Wei, J. (1969). A lumping analysis in monomolecular reaction systems. Analysis of approximaely lumpable system. *Industrial & Engineering chemistry fundamentals* **8**, 124-133.
- Lafrance, P., Lemaire, V., Varin, F., Donati, F., Belair, J., and Nekka, F. (2002). Spatial effects in modeling pharmacokinetics of rapid action drugs. *Bulletin of Mathematical Biology* **64**, 483-500.
- Lazar, N. A., Eddy, W. F., Genovese, C. R., and Welling, J. (2001). Statistical issues in fMRI for brain imaging. *International Statistical Review* **69**, 105-127.
- Le Bihan, D. (1984). Bases Physiques Pour l'Interprétation des Images par Résonance Magnétique Nucléaire. Thèse de la faculté de médecine Pitié-Salpêtrière.
- Li, G., and Rabitz, H. (1989). A general analysis of exact lumping in chemical kinetics. *Chemical Engineering Science* **44**, 1413-1430.
- Li, G., and Rabitz, H. (1990). A general analysis of approximate lumping in chemical kinetics. *Chemical Engineering Science* **45**, 977-1002.
- Li, G., and Rabitz, H. (1991a). Determination of constrained lumping schemes for nonisothermal first-order reaction systems. *Chemical Engineering Science* **46**, 583-596.
- Li, G., and Rabitz, H. (1991b). A general lumping analysis of a reaction system coupled with diffusion. *Chemical Engineering Science* **46**, 2041-2053.
- Li, G., and Rabitz, H. (1991c). New approaches to determination of constrained lumping schemes for a reaction system in the whole composition space. *Chemical Engineering Science* **46**, 95-111.
- Li, G., and Rabitz, H. (1993). The direct lumping approach: an application to a catalytic reforming model. *Chemical Engineering Science* **48**, 1903-1909.
- Li, G., and Rabitz, H. (1997). Reduced kinetic equations of a CO/H<sub>2</sub>/air oxidation model by a special perturbation method. *Chemical Engineering Science* **52**, 4317-4327.
- Li, G., Rabitz, H., and Toth, J. (1994a). A general analysis of exact nonlinear lumping in chemical kinetics. *Chemical Engineering Science* **49**, 343-361.
- Li, G. Y., Tomlin, A. S., Rabitz, H., and Toth, J. (1994b). A general analysis of approximate nonlinear lumping in chemical kinetics. II. Unconstrained lumping. *Journal of Chemical Physics* **101**, 1172-1187.
- Ludden, T. M., Gillespie, W. R., and Bachman, W. J. (1995). Commentary on "Physiologically based pharmacokinetic modeling as a tool for drug development". *Journal of Pharmacokinetics and Biopharmaceutics* **23**, 231-235.
- Ludemann, L., Hamm, B., and Zimmer, C. (2000). Pharmacokinetic analysis of glioma compartments with dynamic Gd-DTPA-enhanced magnetic resonance imaging. *Magn Reson Imaging* **18**, 1201-14.



- Luecke, R. H., Wosilait, W. D., Pearce, B. A., and Young, J. F. (1994). A Physiologically-Based Pharmacokinetic Computer-Model for Human-Pregnancy. *Teratology* **49**, 90-103.
- Lunn, D. J., and Aarons, L. (1998). The pharmacokinetics of saquinavir: A Markov chain Monte Carlo population analysis. *Journal of Pharmacokinetics and Biopharmaceutics* **26**, 47-74.
- Mann, S., Droz, P. O., and Vahter, M. (1996). A physiologically based pharmacokinetic model for arsenic exposure - II. Validation and application in humans. *Toxicology and Applied Pharmacology* **140**, 471-486.
- Margni, M., Rossier, D., Crettaz, P., and Jolliet, O. (2002). Life cycle impact assessment of pesticides on human health and ecosystems. *Agriculture Ecosystems & Environment* **93**, 379-392.
- Mather, L. E. (2001). Anatomical-physiological approaches in pharmacokinetics and pharmacodynamics. *Clinical Pharmacokinetics* **40**, 707-722.
- McGeoch, C. (1992). Analyzing algorithms by simulation: variance reduction techniques and simulation speedups. *ACM Computing Surveys* **24**, 195-212.
- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. (1953). Equation of state calculations by fast computing machines. *Journal of Chemical Physics* **21**, 1087-1092.
- Mezzetti, M., Ibrahim, J. G., Bois, F. Y., Ryan, L. M., Ngo, L., and Smith, T. J. (2003). A Bayesian compartmental model for the evaluation of 1,3-butadiene metabolism. *Journal of the Royal Statistical Society Series C-Applied Statistics* **52**, 291-305.
- Mueller, P. (1999). Simulation Based Optimal Design. In Bayesian Statistics 6 (J. M. B. J.O. Berger, A.P. Dawid and A.F.M. Smith, ed., pp. 459-474. Oxford University Press.
- National Research Council - Committee on the Use of Third Party Toxicity Research with Human Research Participants Science Technology and Law Program (NRC) (2004). *Intentional Human Dosing Studies for EPA Regulatory Purposes: Scientific and Ethical Issues*. The National Academies Press.
- Nestorov, I. (2001). Modelling and simulation of variability and uncertainty in toxicokinetics and pharmacokinetics. *Toxicology Letters* **120**, 411-420.
- Nestorov, I. (2003). Whole body pharmacokinetic models. *Clinical Pharmacokinetics* **42**, 883-908.
- Nestorov, I. A., Aarons, L. J., Arundel, P. A., and Rowland, M. (1998). Lumping of whole-body physiologically based pharmacokinetic models. *Journal of Pharmacokinetics and Biopharmaceutics* **26**, 21-46.
- Okuhata, Y., Brasch, R. C., Pham, C. D., Daldrup, H., Wendland, M. F., Shames, D. M., and Roberts, T. P. (1999). Tumor blood volume assays using contrast-enhanced magnetic resonance imaging: regional heterogeneity and postmortem artifacts. *Journal of Magnetic Resonance Imaging: JMRI* **9**, 685-690.
- Oliver, R. E., Jones, A. F., and Rowland, M. (2001). A whole-body physiologically based pharmacokinetic model incorporating dispersion concepts: Short and long time characteristics. *Journal of Pharmacokinetics and Pharmacodynamics* **28**, 27-55.

Olson, H., Betton, G., Robinson, D., Thomas, K., Monro, A., Kolaja, G., Lilly, P., Sanders, J., Sipes, G., Bracken, W., Dorato, M., Van Deun, K., Smith, P., Berger, B., and Heller, A. (2000). Concordance of the toxicity of pharmaceuticals in humans and in animals. *Regulatory Toxicology and Pharmacology* **32**, 56-67.

Port, M., Corot, C., Raynal, I., Idee, J. M., Dencausse, A., Lancelot, E., Meyer, D., Bonnemain, B., and Lautrou, J. (2001a). Physicochemical and biological evaluation of P792, a rapid- clearance blood-pool agent for magnetic resonance imaging. *Investigative Radiology* **36**, 445-454.

Port, M., Corot, C., Rousseaux, O., Raynal, I., Devoldere, L., Idee, J.-M., Dencausse, A., Greneur, S. L., Simonot, C., and Meyer, D. (2001b). P792: a rapid clearance blood pool agent for magnetic resonance imaging: preliminary results. *Magnetic Resonance Materials in Physics Biology and Medicine* **12**, 121-127.

Port, R. E., Knopp, M. V., and Brix, G. (2001c). Dynamic contrast-enhanced MRI using Gd-DTPA: interindividual variability of the arterial input function and consequences for the assessment of kinetics in tumors. *Magn Reson Med* **45**, 1030-8.

Port, R. E., Knopp, M. V., Hoffmann, U., Milker-Zabel, S., and Brix, G. (1999). Multicompartment analysis of gadolinium chelate kinetics: blood-tissue exchange in mammary tumors as monitored by dynamic MR imaging. *Journal of Magnetic Resonance Imaging* **10**, 233-241.

Poulin, P., Schoenlein, K., and Theil, F. P. (2001). Prediction of adipose tissue : plasma partition coefficients for structurally unrelated drugs. *Journal of Pharmaceutical Sciences* **90**, 436-447.

Poulin, P., and Theil, F. P. (2000). A Priori prediction of tissue : plasma partition coefficients of drugs to facilitate the use of physiologically-based pharmacokinetic models in drug discovery. *Journal of Pharmaceutical Sciences* **89**, 16-35.

Pradel, C., Siauve, N., Bruneteau, G., Clement, O., de Bazelaire, C., Frouin, F., Wedge, S. R., Tessier, J. L., Robert, P. H., Frija, G., and Cuenod, C. A. (2003). Reduced capillary perfusion and permeability in human tumour xenografts treated with the VEGF signalling inhibitor ZD4190: an in vivo assessment using dynamic MR imaging and macromolecular contrast media. *Magnetic Resonance Imaging* **21**, 845-51.

Price, P. S., Conolly, R. B., Chaisson, C. F., Gross, E. A., Young, J. S., Mathis, E. T., and Tedder, D. R. (2003). Modeling interindividual variation in physiological factors used in PBPK models of humans. *Critical Reviews in Toxicology* **33**, 469-503.

Rabe-Hesketh, S., Bullmore, E. T., and Brammer, M. J. (1997). The analysis of functional magnetic resonance images. *Statistical Methods in Medical Research* **6**, 215-37.

Reigner, B. G., Williams, P. E. O., Patel, J. H., Steimer, J. L., Peck, C., and vanBrummelen, P. (1997). An evaluation of the integration of pharmacokinetic and pharmacodynamic principles in clinical drug development - Experience within Hoffmann La Roche. *Clinical Pharmacokinetics* **33**, 142-152.

Robert, C. (1996). *Méthodes de Monte Carlo par Chaînes de Markov*. Economica, Paris.

Rowland, M., Benet, L. Z., and Graham, G. G. (1973). Clearance concepts in pharmacokinetics. *Journal of Pharmacokinetics and Biopharmaceutics* **1**, 123-136.

Rowland, M., and Tozer, T. N. (1989). *Clinical Pharmacokinetics: Concepts and Applications*. Lea & Febiger, Philadelphia.

Rudin, M., Beckmann, N., Porszasz, R., Reese, T., Bochelen, D., and Sauter, A. (1999). In vivo magnetic resonance methods in pharmaceutical research: current status and perspectives. *Nmr in Biomedicine* **12**, 69-97.

Sato, A., Nakajima, T., Fujiwara, Y., and Murayama, N. (1977). A pharmacokinetic model to study the excretion of trichloroethylene and its metabolites after an inhalation exposure. *British Journal of Industrial Medicine* **34**, 56-63.

Sheiner, L. B., Beal, S., and Weichman, B. UCSF NONMEM software. *licensed and distributed by GloboMax LLC*, <http://c255.ucsf.edu/nonmem0.html>.

Sheiner, L. B., and Beal, S. L. (1980). Evaluation of methods for estimating population pharmacokinetic parameters - I. Michaelis-Menten model: routine clinical pharmacokinetic data. *Journal of Pharmacokinetics and Biopharmaceutics* **8**, 553-571.

Sheiner, L. B., and Ludden, T. M. (1992). Population pharmacokinetics/dynamics. *Annual Review of Pharmacology and Toxicology* **32**, 185-209.

Simmons, J. E. (1996). Application of physiologically based pharmacokinetic modelling to combination toxicology. *Food and Chemical Toxicology* **34**, 1067-1073.

Singer, M. A. (2001). Of mice and men and elephants: metabolic rate sets glomerular filtration rate. *American Journal of Kidney Diseases* **37**, 164-178.

Smith, T., Lin, Y.-S., Mezzetti, L., Bois, F. Y., Kelsey, K., and Ibrahim, J. (2001). Genetic and dietary factors affecting human metabolism of 1,3-butadiene. *Chemico-Biological Interactions* **135-136**, 407-428.

Spiegelhalter, D., Thomas, A., Best, N., and Gilks, W. R. (1996). Bayesian inference Using Gibbs Sampling - Manual. Institute of Public Health, MRC Biostatistics Unit, Cambridge.

Suzuki, H., Iwatsubo, T., and Sugiyama, Y. (1995). Applications and prospects for physiologically based pharmacokinetic (PB-PK) models involving pharmaceutical agents. *Toxicology Letters* **82-83**, 349-55.

Sweeney, L. M., Tyler, T. R., Kirman, C. R., Corley, R. A., Reitz, R. H., Paustenbach, D. J., Holson, J. F., Whorton, M. D., Thompson, K. M., and Gargas, M. L. (2001). Proposed occupational exposure limits for select ethylene glycol ethers using PBPK models and Monte Carlo simulations. *Toxicological Sciences* **62**, 124-139.

Taylor, J. S., and Reddick, W. E. (2000). Evolution from empirical dynamic contrast-enhanced magnetic resonance imaging to pharmacokinetic MRI. *Advanced Drug Delivery Reviews* **41**, 91-110.

Taylor, J. S., Tofts, P. S., Port, R., Evelhoch, J. L., Knopp, M., Reddick, W. E., Runge, V. M., and Mayr, N. (1999). MR imaging of tumor microcirculation: promise for the new millennium. *Journal of Magnetic Resonance Imaging* **10**, 903-907.

Teorell, T. (1937). Kinetics of distribution of substances administered to the body. *Archives Internationales de Pharmacodynamie et de Thérapie* **57**, 205-240.

Theil, F. P., Guentert, T. W., Haddad, S., and Poulin, P. (2003). Utility of physiologically based pharmacokinetic models to drug development and rational drug discovery candidate selection. *Toxicology Letters* **138**, 29-49.

Thomas, A., Spiegelhalter, D. J., and Gilks, W. R. (1992). BUGS: a program to perform Bayesian inference using Gibbs sampling. In *Bayesian Statistics 4* (A. F. M. Smith, ed., pp. 837-842. Oxford University Press, Oxford.

Tierney, L., and Mira, A. (1999). Some adaptive Monte Carlo methods for Bayesian inference. *Statistics in Medicine* **18**, 2507-2515.

Timchalk, C., Walker, N. J., Mann, R. C., and Metting, F. B. (2001). The virtual body workshop: current and future application of human biology models in environmental health research. *Environ Health Perspect* **109**, 421-3.

Tofts, P. S. (1997). Modeling tracer kinetics in dynamic Gd-DTPA MR imaging. *Journal of Magnetic Resonance Imaging* **7**, 91-101.

Tofts, P. S., Berkowitz, B., and Schnall, M. D. (1995). Quantitative analysis of dynamic Gd-DTPA enhancement in breast tumors using a permeability model. *Magnetic Resonance in Medicine* **33**, 564-8.

Tofts, P. S., and Berkowitz, B. A. (1994). Measurement of capillary permeability from the Gd enhancement curve: a comparison of bolus and constant infusion injection methods. *Magnetic Resonance Imaging* **12**, 81-91.

Tofts, P. S., Brix, G., Buckley, D. L., Evelhoch, J. L., Henderson, E., Knopp, M. V., Larsson, H. B., Lee, T. Y., Mayr, N. A., Parker, G. J., Port, R. E., Taylor, J., and Weisskoff, R. M. (1999). Estimating kinetic parameters from dynamic contrast-enhanced T(1)-weighted MRI of a diffusable tracer: standardized quantities and symbols. *Journal of Magnetic Resonance Imaging* **10**, 223-32.

Tofts, P. S., and Kermode, A. G. (1991). Measurement of the blood-brain barrier permeability and leakage space using dynamic MR imaging. 1. Fundamental concepts. *Magnetic Resonance in Medicine* **17**, 357-67.

Tomlin, A. S., Turanyi, T., and Pilling, M. J. (1997). Mathematical tools for the construction, investigation and reduction of combustion mechanisms. In *Low-temperature combustion and autoignition* (M. J. Pilling, ed., Vol. 35, pp. 293-437. Elsevier, Amsterdam.

Toth, J., Li, G. Y., Rabitz, H., and Tomlin, A. S. (1997). Effect of lumping and expanding on kinetic differential equations. *Siam Journal on Applied Mathematics* **57**, 1531-1556.

Tozer, T. N. (1981). Concepts basic to pharmacokinetics. *Pharmacology and Therapeutics* **12**, 109-131.

US Department of Labor Occupational Safety and Health Administration (1994). Sampling and Analytical Methods, Enflurane, Halothane, Isoflurane, Method number: 103. <http://www.osha.gov/dts/sltc/methods/index.html>.

US Environmental Protection Agency (1996). Proposed guidelines for carcinogenic risk assesement. EPA, Washington, DC.

US Environmental Protection Agency (2000). Comments on the use of data from the testing of human subjects. Science Advisory Board, Washington, DC.

- Van der Mollen, G. W., Kooijman, S. A. L. M., and Slob, W. (1996). A generic toxicokinetic model for persistent lipophilic compounds in humans: an application to TCDD. *Fundamental and Applied Toxicology* **31**, 83-94.
- Veng-Pedersen, P., and Modi, N. B. (1992). Neural networks in pharmacodynamic modeling. Is current modeling practice of complex kinetic systems at a dead end? *Journal of Pharmacokinetics and Biopharmaceutics* **20**, 397-412; discussion 413-8.
- Verhaar, H. J., Morroni, J. R., Reardon, K. F., Hays, S. M., Gaver, D. P., Jr., Carpenter, R. L., and Yang, R. S. (1997). A proposed approach to study the toxicology of complex mixtures of petroleum products: the integrated use of QSAR, lumping analysis and PBPK/PD modeling. *Environmental Health Perspectives* **105 Suppl 1**, 179-95.
- Vinegar, A., Jepson, G. W., and Overton, J. H. (1998). PBPK modeling of short-term (0 to 5 min) human inhalation exposures to halogenated hydrocarbons. *Inhalation Toxicology* **10**, 411-429.
- Wakefield, J. (2000). The combination of population pharmacokinetic studies. *Biometrics* **56**, 263-270.
- Wakefield, J., and Bennett, J. (1996). The Bayesian modeling of covariates for population pharmacokinetic models. *Journal of the American Statistical Association* **91**, 917-927.
- Wakefield, J. C., and Racine-Poon, A. (1995). An application of Bayesian population pharmacokinetic pharmacodynamic models to dose recommendation. *Statistics in Medicine* **14**, 971-986.
- Wang, S. W., Georgopoulos, P. G., Li, G., and Rabitz, H. (1998). Condensing complex atmospheric chemistry mechanisms. 1. The direct constrained approximate lumping (DCAL) method applied to alkane photochemistry. *Environmental Science & Technology* **32**, 2018-2024.
- Ward, R. C., Yambert, M. W., Toedte, R. J., Munro, N. B., Easterly, C. E., Difilippo, E. P., and Stallings, D. C. (2000). Creating a human phantom for the virtual human program. *Studies in Health Technology and Informatics* **70**, 368-74.
- Watanabe, K., and Bois, F. Y. (1996). Interspecies extrapolation of physiological pharmacokinetic parameter distributions. *Risk Analysis* **16**, 741-754.
- Wedge, S. R., Ogilvie, D. J., Dukes, M., Kendrew, J., Curwen, J. O., Hennequin, L. F., Thomas, A. P., Stokes, E. S. E., Curry, B., Richmond, G. H. P., and Wadsworth, P. F. (2000). ZD4190: An orally active inhibitor of vascular endothelial growth factor signaling with broad-spectrum antitumor efficacy. *Cancer Research* **60**, 970-975.
- Wei, J., and Kuo, J. C. W. (1969). A lumping analysis in monomolecular reaction systems. Analysis of the exactly lumpable system. *Industrial & Engineering chemistry fundamentals* **8**, 114-123.
- Welsch, F., Blumenthal, G. M., and Conolly, R. B. (1995). Physiologically based pharmacokinetic models applicable to organogenesis: extrapolation between species and potential use in prenatal toxicity risk assessments. *Toxicology Letters* **82-83**, 539-547.
- West, G. B., Brown, J. H., and Enquist, B. J. (1997). A general model for the origin of allometric scaling laws in biology. *Science* **276**, 122-126.

Zeise, L., Hattis, D., Andersen, M., Bailer, A. J., Bayard, S., Chen, C., Clewell, H., Conolly, R., Crump, K., Dunson, D., Finkel, A., Haber, L., Jarabek, A. M., Kodell, R., Krewski, D., Thomas, D., Thorslund, T., and Wassell, J. T. (2002). Improving risk assessment: Research opportunities in dose response modeling to improve risk assessment. *Human and Ecological Risk Assessment* **8**, 1421-1444.

## 10. Publications, communications, et collaborations

### 10.1. Publications

C. Brochot, D. Balvay, B. Bessoud, N. Siauve, CA. Cuenod, FY. Bois. *Bayesian physiologically based pharmacokinetic modelling of tumor's microcirculation magnetic resonance imaging data : Application to the evaluation of an antiangiogenic treatment*. Soumis (en révision) à Magnetic Resonance in Medicine.

C. Brochot, FY. Bois. *Use of a chemical probe to increase safety for human volunteers in toxicokinetic studies*. Soumis à Risk Analysis.

S. Micallef, C. Brochot, FY. Bois. *L'analyse statistique bayésienne de données toxicocinétiques*. Soumis à Environnement, Risques et Santé.

C. Brochot, J. Tóth, FY. Bois. *Lumping in toxicokinetics*. En préparation.

### 10.2. Communications

*Reducing the risk incurred by human volunteers in toxicokinetic studies through the use of tracers*, C. Brochot, FY. Bois. Communication écrite au congrès Exposure assessment in a changing environment (X2004), juin 2004, Utrecht, Pays-Bas.

*Bayesian physiologically based pharmacokinetic modelling of tumor's microcirculation magnetic resonance imaging data: Application to the evaluation of an antiangiogenic treatment*, C. Brochot, D. Balvay, B. Bessoud, N. Siauve, CA. Cuenod, FY. Bois. Communication écrite au congrès In silico driven drug discovery and development (Cost B15), mai 2004, Genève, Suisse.

*Lumping on Toxicokinetics*, J. Tóth, C. Brochot, FY. Bois. Communication écrite au congrès European Science Foundation Reactor Workshop, janvier 2003, Budapest, Hongrie.

*Analyse d'images biomédicales : Application à la perméabilité vasculaire des tumeurs*, C. Brochot, FY. Bois. Communication orale au séminaire de recherche "Statistiques spatiales" à l'Institut National Agronomique de Paris-Grignon, mai 2002, Paris, France.

*Application of aggregation techniques to toxicokinetic modelling of toxic substances*, C. Brochot, FY. Bois. Communication orale lors d'un séminaire de mathématiques organisé par le Dr. J. Tóth, Département de Mathématiques, Université de Technologie et Sciences Economiques, mars 2002, Budapest, Hongrie.

### 10.3. Collaborations

Pour mener le projet concernant les données d'imagerie, une collaboration a été montée avec l'unité INSERM U494 "Imagerie Médicale Quantitative", et plus particulièrement l'équipe "Biodistribution de marqueurs, applications aux tumeurs solides". Au cours de ce projet, notre principal interlocuteur a été le Dr. Charles-André Cuenod.

Le travail sur l'utilisation de traceurs non-toxiques fait parti d'un projet géré par l'Ecole de Santé Publique de l'Université de Harvard, Etats-Unis. Les études sur la toxicocinétique du 1,3-butadiène sont menées en collaboration avec l'équipe du Pr. Thomas J. Smith.

La partie d'agrégation des systèmes d'équations différentielles a été réalisée en collaboration avec le Pr. J. Tóth, appartenant au département de mathématiques de l'Université de Technologie et Sciences Economiques de Budapest, Hongrie.



# RESUME

Les modèles physiologiques toxicocinétiques (ou PBPK pour Physiologically Based Pharmacokinetic) permettent de décrire la biodistribution d'une substance (c'est-à-dire, son administration, distribution, métabolisme, et excrétion) au sein d'un organisme. Cependant, l'ajustement d'un modèle PBPK à des données expérimentales peut se révéler difficile en raison de la complexité du modèle, du nombre important de paramètres à estimer, et de la nature des données toxicocinétiques (généralement peu nombreuses, éparpillées et entachées d'erreur de mesure). Pour pallier à ces difficultés, nous proposons de développer deux protocoles expérimentaux facilitant la paramétrisation des modèles PBPK, par l'apport d'informations complémentaires sur les paramètres des modèles : la collecte de données sur la distribution spatio-temporelle d'une substance dans l'organisme par des techniques d'imagerie, et l'utilisation de traceurs non-toxiques lors d'études toxicocinétiques.

Pour notre premier protocole, nous avons appliqué l'analyse d'image au développement d'un modèle PBPK décrivant la biodistribution d'un agent de contraste, Vistarem®. Grâce à la modélisation des données recueillies par imagerie par résonance magnétique, nous avons pu caractériser la vascularisation du corps entier et celle d'une tumeur greffée, ainsi qu'évaluer statistiquement les effets d'un traitement antiangiogénique.

Ensuite, nous avons introduit l'utilisation de traceurs explorateurs lors d'études toxicocinétiques. Nous avons développé des protocoles d'exposition à une substance toxique et à un traceur non-toxique pour les humains. Sur la base d'expositions simulées, nous avons montré qu'en administrant aux sujets deux substances, il est possible de réduire la dose d'exposition du toxique (par 4 dans notre exemple d'application) tout en maintenant la précision sur l'estimation des paramètres d'intérêt. L'inconvénient majeur de ce protocole d'exposition est la lourdeur des calculs induits (la calibration de deux modèles PBPK est nécessaire). Afin de réduire la complexité des calculs engendrés, nous avons proposé d'appliquer des techniques d'agrégation de systèmes d'équations différentielles. L'essentiel de ces techniques est de réduire un système tout en conservant sa dynamique. Sur des exemples simples, nous avons montré leur potentiel en toxico/pharmacocinétique.

Chacun des deux protocoles développés dans ce travail inclut, en plus de la substance d'intérêt, l'utilisation d'une substance "annexe". Notre travail montre que l'utilisation de cette substance annexe permet de renforcer la connaissance sur l'anatomie et la physiologie du sujet considéré. Ceci conduit donc à une meilleure détermination de l'action de la substance d'intérêt. Une suite logique de notre travail serait de coupler ces deux protocoles, c'est-à-dire de développer des protocoles d'exposition à une substance toxique et à une substance visualisable par une technique d'imagerie quantitative.

**Mots-Clés :** modèles PBPK; analyse statistique bayésienne; imagerie par résonance magnétique; traceur; méthode d'agrégation.